PDGFRA and CCNB1 are promising druggable targets for treating Ovarian Neoplasms that control activity of EP300, NFYA and HSF1 transcription factors on of differentially expressed genes

Demo User
geneXplain GmbH
info@genexplain.com
Data received on 26/11/2021 ; Run on 02/12/2024 ; Report generated on 02/12/2024

Genome Enhancer release 3.5 (TRANSFAC®, TRANSPATH® and HumanPSD™ release 2024.2)

Abstract

In the present study we applied the software package "Genome Enhancer" to a multiomics data set that contains *transcriptomics and epigenomics* data. The study is done in the context of *Ovarian Neoplasms*. The goal of this pipeline is to identify potential drug targets in the molecular network that governs the studied pathological process. In the first step of analysis pipeline discovers transcription factors (TFs) that regulate genes activities in the pathological state. The activities of these TFs are controlled by so-called master regulators, which are identified in the second step of analysis. After a subsequent druggability checkup, the most promising master regulators are chosen as potential drug targets for the analyzed pathology. At the end the pipeline comes up with (a) a list of known drugs and (b) investigational active chemical compounds with the potential to interact with selected drug targets.

From the data set analyzed in this study, we found the following TFs to be potentially involved in the regulation of the differentially expressed genes: EP300, NFYA, RXRA, HSF1 and RELA. The subsequent network analysis suggested

- PDGFRalpha
- vrk1
- Cdk1-isoform1(h):cyclinB1-isoform1

as the most promising molecular targets for further research, drug development and drug repurposing initiatives on the basis of identified molecular mechanism of the studied pathology. Having checked the actual druggability potential of the full list of identified targets, both, via information available in medical literature and via cheminformatics analysis of drug compounds, we have identified the following drugs as the most promising treatment candidates for the studied pathology: Pazopanib, fimepinostat, Paclitaxel and Etoposide.

1. Introduction

Recording "-omics" data to measure gene activities, protein expression or metabolic events is becoming a standard approach to characterize the pathological state of an affected organism or tissue. Increasingly, several of these methods are applied in a combined approach leading to large "multiomics" datasets. Still the challenge remains how to reveal the underlying molecular mechanisms that render a given pathological state different from the norm. The disease-causing mechanism can be described by a re-wiring of the cellular regulatory network, for instance as a result of a genetic or epigenetic alterations influencing the activity of relevant genes. Reconstruction of the disease-specific regulatory networks can help identify potential master regulators of the respective pathological process. Knowledge about these master regulators can point to ways how to block a pathological regulatory cascade. Suppression of certain molecular targets as components of these cascades may stop the pathological process and cure the disease.

Conventional approaches of statistical "-omics" data analysis provide only very limited information about the causes of the observed phenomena and therefore contribute little to the understanding of the pathological molecular mechanism. In contrast, the "upstream analysis" method [1-4] applied here has been deviced to provide a casual interpretation of the data obtained for a pathology state. This approach comprises two major steps: (1) analysing promoters and enhancers of differentially expressed genes for the transcription factors

(TFs) involved in their regulation and, thus, important for the process under study; (2) re-constructing the signaling pathways that activate these TFs and identifying master regulators at the top of such pathways. For the first step, the database TRANSFAC® [6] is employed together with the TF binding site identification algorithms Match [7] and CMA [8]. The second step involves the signal transduction database TRANSPATH® [9] and special graph search algorithms [10-11] implemented in the software "Genome Enhancer".

The "upstream analysis" approach has now been extended by a third step that reveals known drugs suitable to inhibit (or activate) the identified molecular targets in the context of the disease under study. This step is performed by using information from HumanPSD™ database [5]. In addition, some known drugs and investigational active chemical compounds are subsequently predicted as potential ligands for the revealed molecular targets. They are predicted using a pre-computed database of spectra of biological activities of chemical compounds of a library of 2245 known drugs and investigational chemical compounds from HumanPSD™ database. The spectra of biological activities for these compounds are computed using the program PASS on the basis of a (Q)SAR approach [12-14]. These predictions can be used for the research purposes - for further drug development and drug repurposing initiatives.

2. Data

For this study the following experimental data was used:

Table 1. Experimental datasets used in the study

File name	Data type
GSM385721.CEL	Transcriptomics
GSM385722.CEL	Transcriptomics
GSM385723.CEL	Transcriptomics
GSM385724.CEL	Transcriptomics
GSM385725.CEL	Transcriptomics
GSM385726.CEL	Transcriptomics
GSM385727.CEL	Transcriptomics
GSM385728.CEL	Transcriptomics
GSM385729.CEL	Transcriptomics
GSM385730.CEL	Transcriptomics
GSM385747_CpG_NM.fixed.hg38.top300	Epigenomics

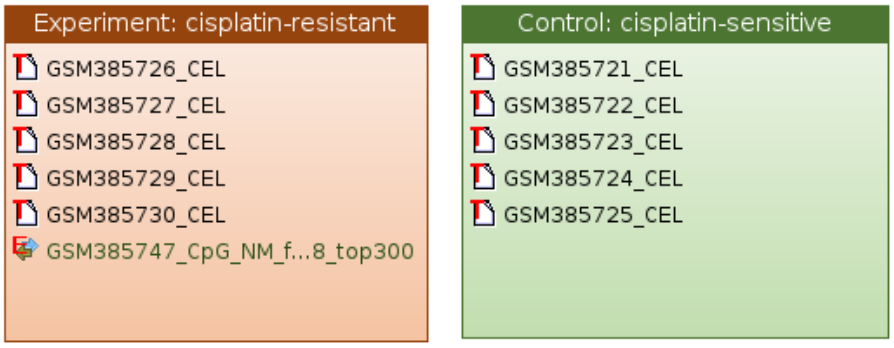


Figure 1. Annotation diagram of experimental data used in this study. With the colored boxes we show those sub-categories of the data that are compared in our analysis.

3. Results

We have compared the following conditions: Experiment: cisplatin-resistant *versus* Control: cisplatin-sensitive.

3.1. Identification of target genes

In the first step of the analysis ***target genes*** were identified from the uploaded experimental data. We applied the Limma tool (R/Bioconductor package integrated into our pipeline) and compared gene expression in the following sets: "Experiment: cisplatin-resistant" with "Control: cisplatin-sensitive". Limma calculated the LogFC (the logarithm to the base 2 of the fold change between different conditions), the p-value and the adjusted p-value (corrected for multiple testing) of the observed fold change. As a result, we detected 4095 upregulated genes (LogFC>0.1) out of which 3369 genes were found as significantly upregulated (p-value<0.1) and 4204 downregulated genes (LogFC<-0.1) out of which 3383 genes were significantly downregulated (p-value<0.1). See tables below for the top significantly up- and downregulated genes. Below we call **target genes** the full list of up- and downregulated genes revealed in our analysis (see tables in Supplementary section).

*Table 2. Top ten significant **up-regulated** genes in Experiment: cisplatin-resistant vs. Control: cisplatin-sensitive.*
See full table →

ID	Gene symbol	Gene description	logFC	P.Value	adj.P.Val
ENSG00000123700	KCNJ2	potassium inwardly rectifying channel subfamily J member 2	5.37	8.03E-14	8.68E-11
ENSG00000064218	DMRT3	doublesex and mab-3 related transcription factor 3	4.03	7.06E-12	2.46E-9
ENSG00000099139	PCSK5	proprotein convertase subtilisin/kexin type 5	3.93	1.32E-14	2.03E-11
ENSG00000197705	KLHL14	kelch like family member 14	3.89	9.95E-13	4.31E-10
ENSG00000129038	LOXL1	lysyl oxidase like 1	3.54	2.19E-10	3.16E-8
ENSG00000133083	DCLK1	doublecortin like kinase 1	3.24	7.37E-13	3.62E-10
ENSG00000141431	ASXL3	ASXL transcriptional regulator 3	3.14	1.34E-11	3.63E-9
ENSG00000126950	TMEM35A	transmembrane protein 35A	3.05	1.47E-12	5.88E-10
ENSG00000164692	COL1A2	collagen type I alpha 2 chain	2.86	2.16E-10	3.15E-8
ENSG00000138378	STAT4	signal transducer and activator of transcription 4	2.86	2.95E-10	3.72E-8

*Table 3. Top ten significant **down-regulated** genes in Experiment: cisplatin-resistant vs. Control: cisplatin-sensitive.*
See full table →

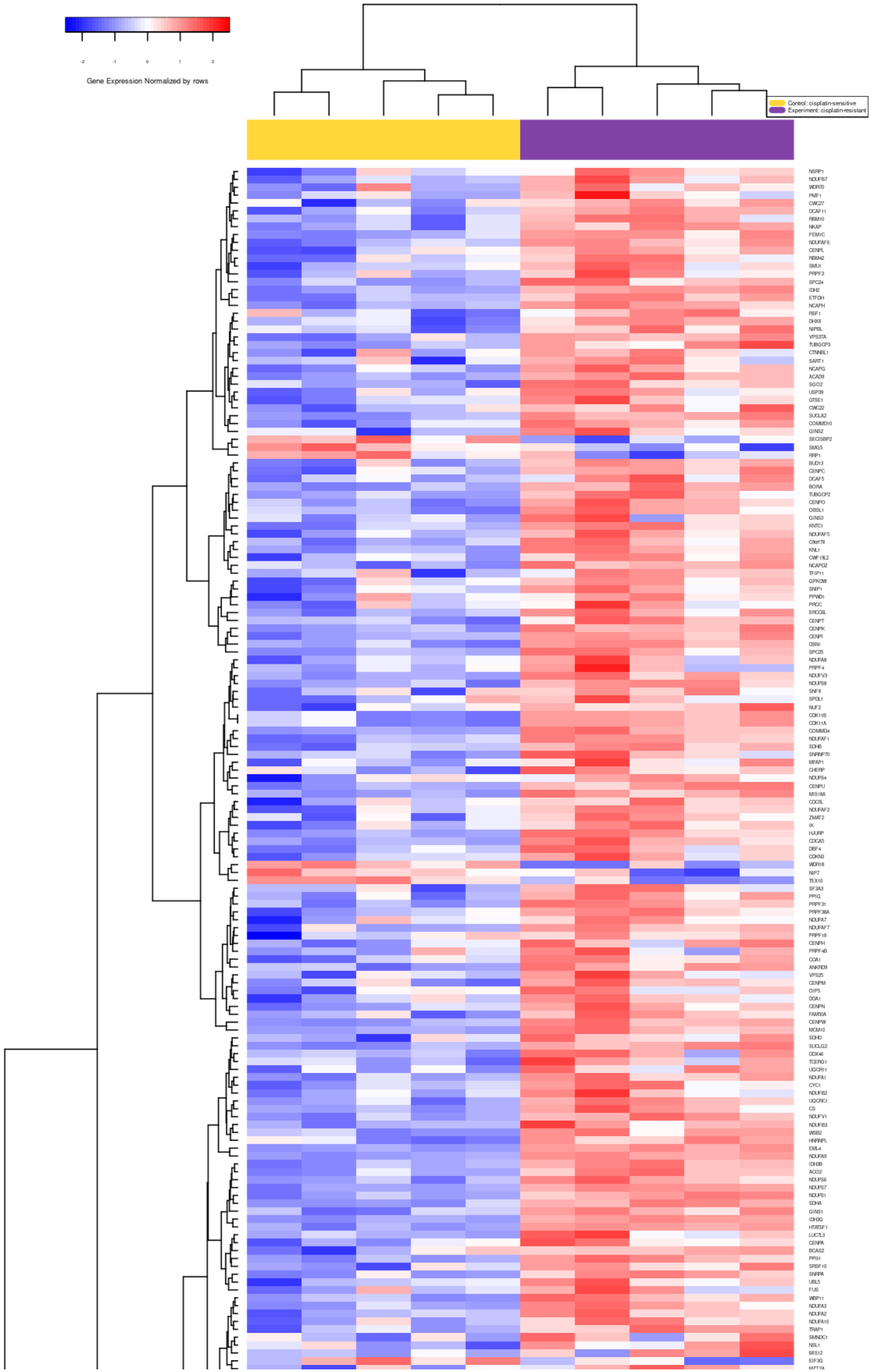
ID	Gene symbol	Gene description	logFC	P.Value	adj.P.Val
ENSG00000127324	TSPAN8	tetraspanin 8	-6.39	1.45E-15	3.93E-12
ENSG00000139292	LGR5	leucine rich repeat containing G protein-coupled receptor 5	-6.25	5.63E-18	6.09E-14
ENSG00000149968	MMP3	matrix metallopeptidase 3	-5.16	1.88E-13	1.56E-10
ENSG00000163359	COL6A3	collagen type VI alpha 3 chain	-5.08	4.57E-16	1.65E-12
ENSG00000169908	TM4SF1	transmembrane 4 L six family member 1	-4.94	1.51E-16	8.17E-13
ENSG00000153233	PTPRR	protein tyrosine phosphatase receptor type R	-4.6	5.9E-13	3.19E-10
ENSG00000166670	MMP10	matrix metallopeptidase 10	-4.45	8.22E-15	1.48E-11
ENSG00000106511	MEOX2	mesenchyme homeobox 2	-4.26	3.59E-12	1.34E-9
ENSG00000145431	PDGFC	platelet derived growth factor C	-4.15	3.14E-14	4.25E-11
ENSG00000060718	COL11A1	collagen type XI alpha 1 chain	-3.65	9.26E-11	1.7E-8

3.2. Functional classification of genes

A functional analysis of differentially expressed genes was done by mapping the significant up-regulated and significant down-regulated genes to several known ontologies, such as Gene Ontology (GO), disease ontology (based on HumanPSD™ database) and the ontology of signal transduction and metabolic pathways from the TRANSPATH® database. Statistical significance was computed using a binomial test.
Figures 3-8 show the most significant categories.

Heatmap of differentially expressed genes in Experiment: cisplatin-resistant vs. Control: cisplatin-sensitive

A heatmap of all differentially expressed genes playing a potential regulatory role in the system (enriched in TRANSPATH® pathways) is presented in Figure 2.


-2
-1
0
1
2
Gene Expression Normalized by rows
Control: cisplatin-sensitive
Experiment: cisplatin-resistant
NSRP1
NDUFB7
WDR70
PMF1
CWC27
DCAF11
RBM10
NKAP
FEMUC
NDUFAF6
CENPL
RBM42
SMU1
PRPF3
SPC24
IDH2
ETFDH
NCAPH
RSF1
DHX8
NIPBL
VPS37A
TUBGCP3
CTNNBL1
SART1
NCAPG
ACAD9
SGO2
USP39
GTSE1
CWC22
SUCLA2
COMMD10
GINS2
SECISBP2
SMG5
RPP1
BUD13
CENPC
DCAF5
BORA
TUBGCP2
CENPO
CIBSL1
GINS3
KNTC1
NDUFAF5
C9orf78
KNL1
CWF19L2
NCAPD2
TFIP11
GPKOW
SNIP1
PPWD1
PROC
ERCC6L
CENPT
CENPK
CENPI
DSN1
SPC25
NDUFA8
PRPF4
NDUFV3
NDUFS8
SNF8
SPDL1
NUF2
CDK11B
CDK11A
COMMD4
NDUFAF1
SDHB
SNRNP70
MFAP1
CHERP
NDUFS4
CENPU
MIS18A
CDC5L
NDUFAF2
ZMAT2
IK
HJURP
CDCA5
DBF4
CDKN3
WDR18
NIP7
TEX10
SF3A3
PPIG
PRPF31
PRPF38A
NDUFA7
NDUFAF7
PRPF18
CENPH
PRPF4B
COA1
ANKRD9
VPS25
CENPM
OIP5
DDA1
CENPN
FAM50A
CENPW
MCM10
SDHD
SUCLG2
DDX46
TCERG1
UQCR11
NDUFA1
CYC1
NDUFB2
UQCRC1
CS
NDUFV1
NDUFB3
WSB2
HNRNPL
EML4
NDUFA9
IDH3B
ACO2
NDUFS6
NDUFS7
NDUFS1
SDHA
GINS1
IDH3G
HTATSF1
LUC7L3
CENPA
BCAS2
PPIH
SRSF10
SNRPA
UBL5
FUS
WBP11
NDUFA3
NDUFA2
NDUFA10
TRAP1
SMNDC1
NSL1
MIS12
EIF3G
MTTP

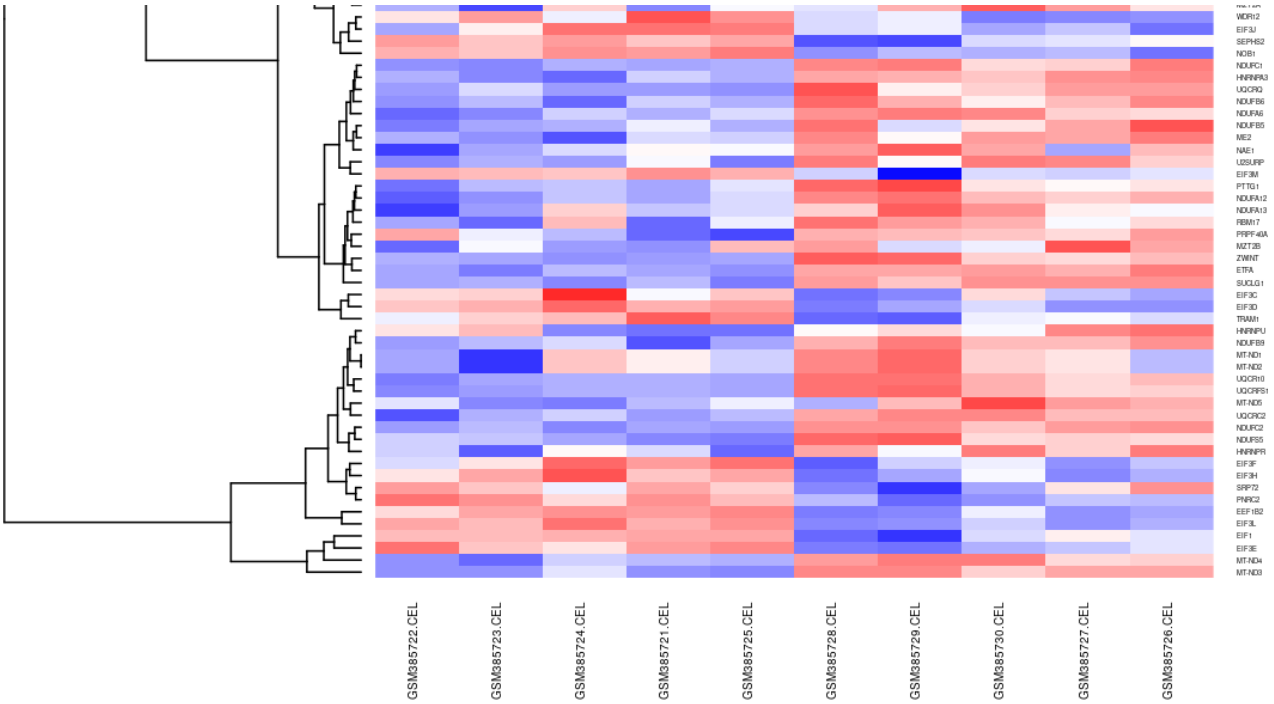


Figure 2. Heatmap of genes enriched in Transpath categories. The colored bar at the top shows the types of the samples according to the legend in the upper right corner.

See full diagram →

Up-regulated genes in Experiment: cisplatin-resistant vs. Control: cisplatin-sensitive:

3369 significant up-regulated genes were taken for the mapping.

GO (biological process)

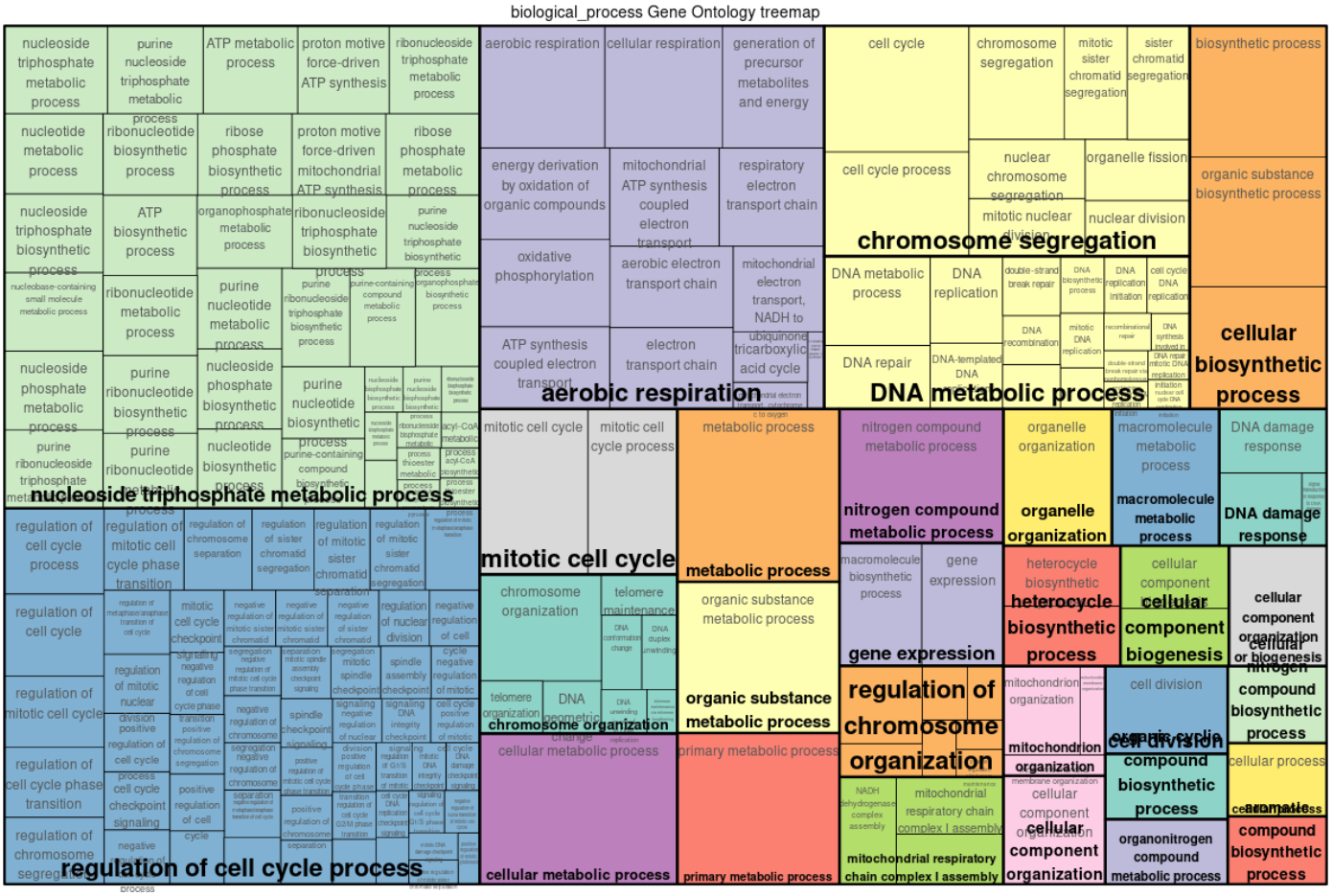


Figure 3. Enriched GO (biological process) of up-regulated genes in Experiment: cisplatin-resistant vs. Control: cisplatin-sensitive.

Full classification →

TRANSPATH® Pathways (2024.2)

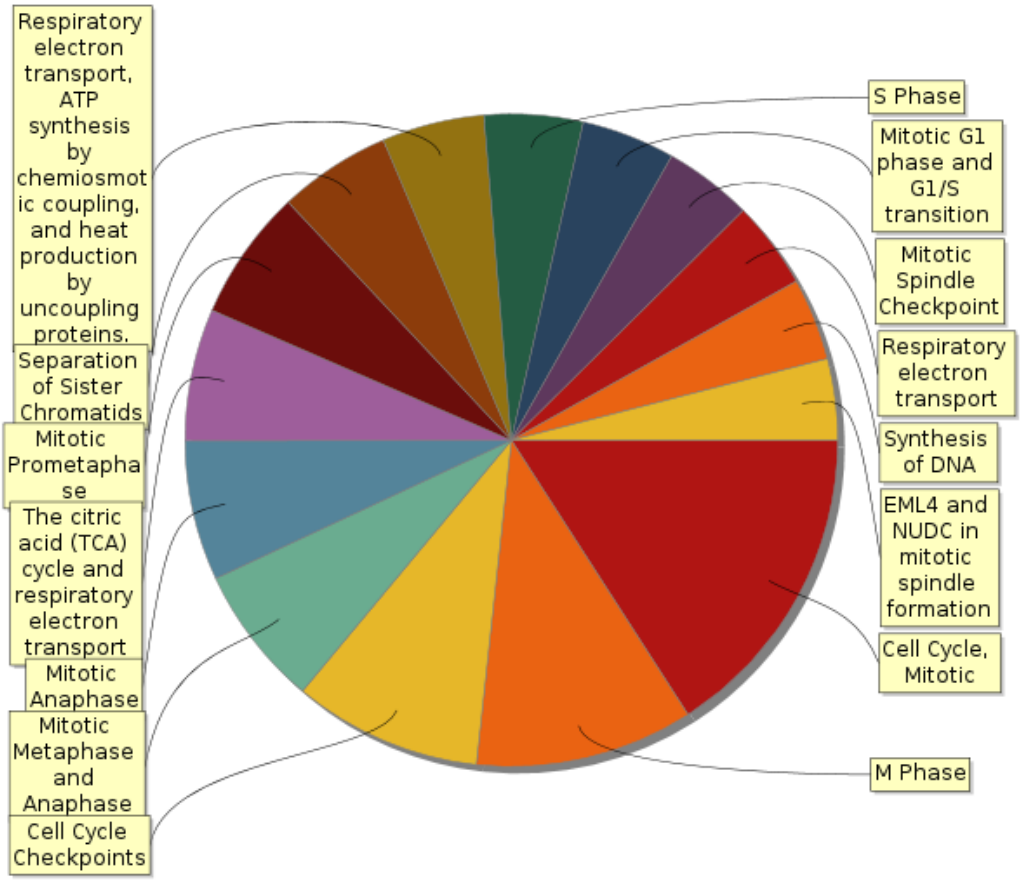


Figure 4. Enriched TRANSPATH® Pathways (2024.2) of up-regulated genes in Experiment: cisplatin-resistant vs. Control: cisplatin-sensitive.
Full classification →

HumanPSD(TM) disease (2024.2)

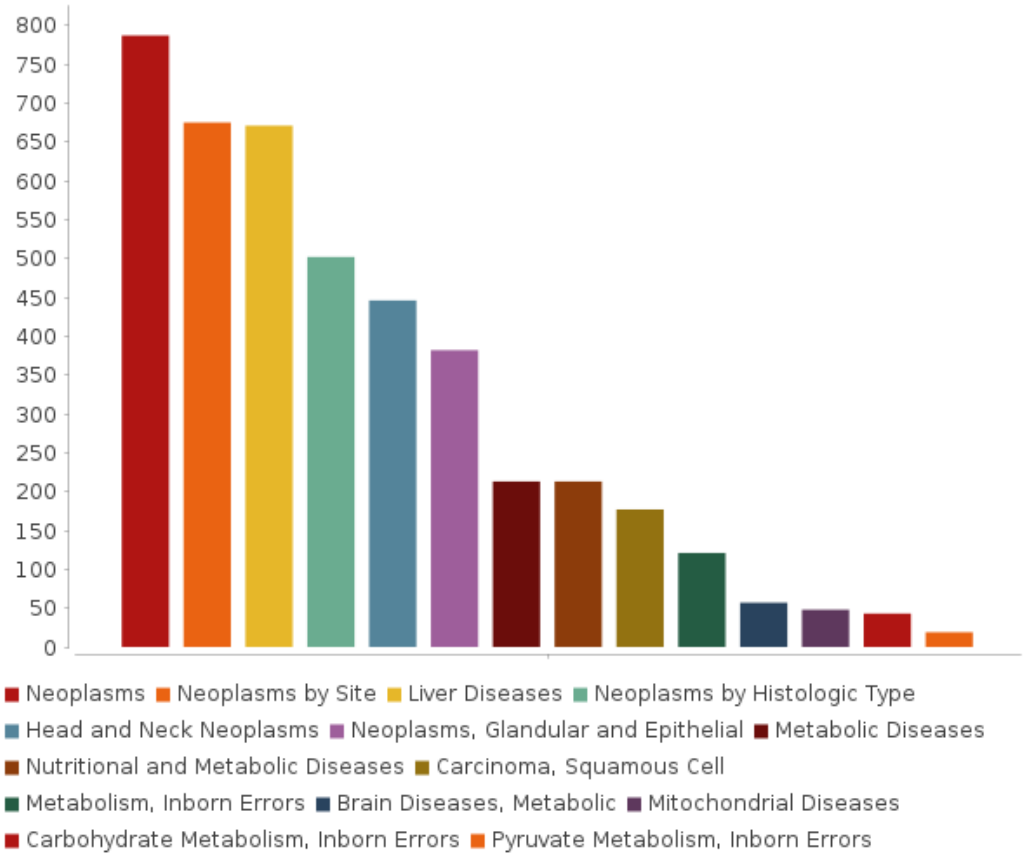


Figure 5. Enriched HumanPSD(TM) disease (2024.2) of up-regulated genes in Experiment: cisplatin-resistant vs. Control: cisplatin-sensitive. The size of the bars correspond to the number of biomarkers of the given disease found among the input set.

Full classification →

Down-regulated genes in Experiment: cisplatin-resistant vs. Control: cisplatin-sensitive:

3383 significant down-regulated genes were taken for the mapping.

GO (biological process)

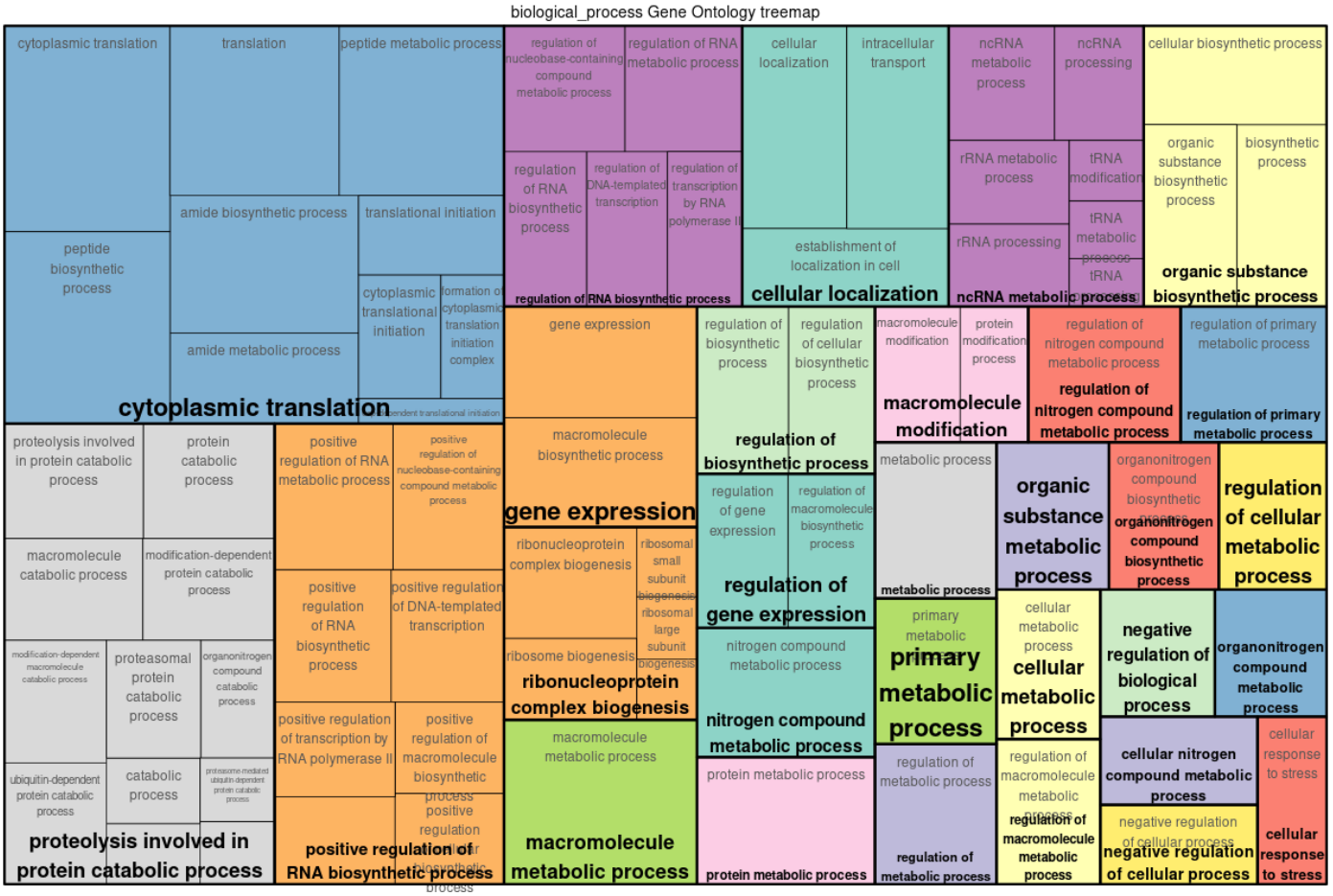


Figure 6. Enriched GO (biological process) of down-regulated genes in Experiment: cisplatin-resistant vs. Control: cisplatin-sensitive.

Full classification →

TRANSPATH® Pathways (2024.2)

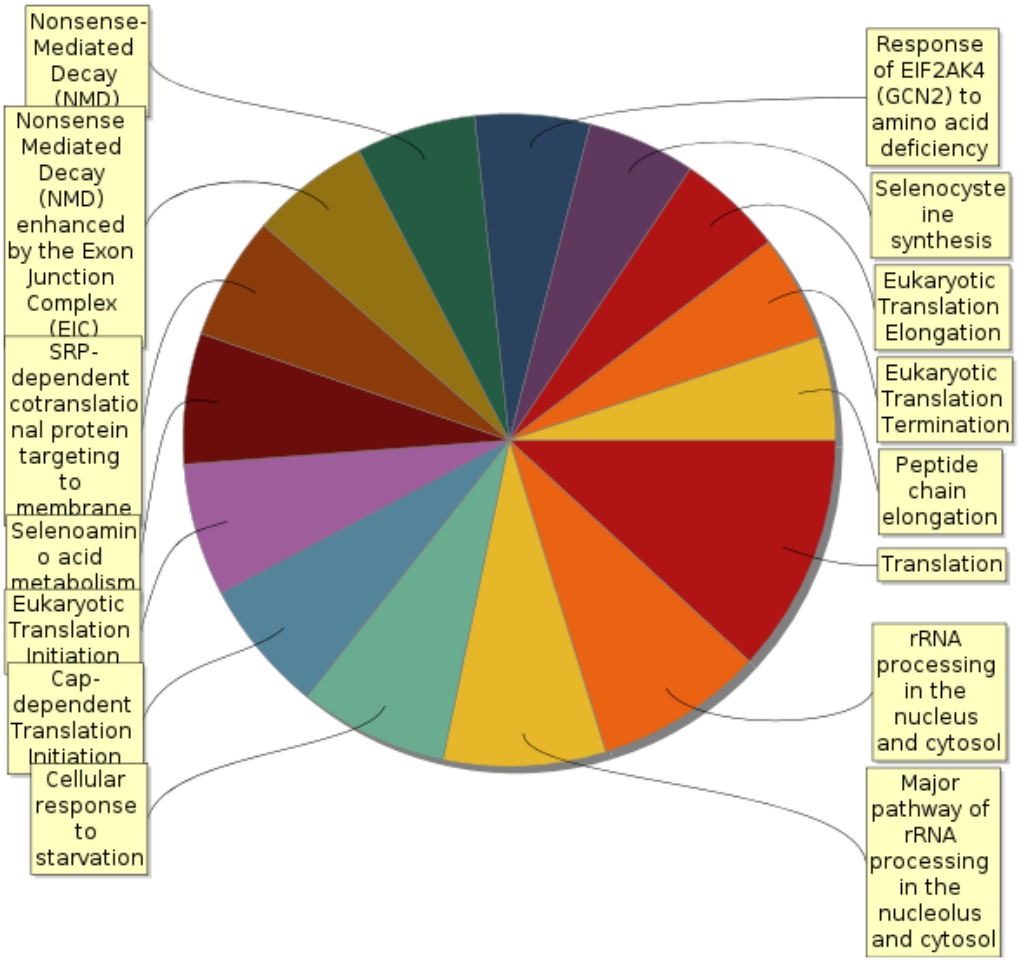


Figure 7. Enriched TRANSPATH® Pathways (2024.2) of down-regulated genes in Experiment: cisplatin-resistant vs. Control: cisplatin-sensitive.
Full classification →

HumanPSD(TM) disease (2024.2)

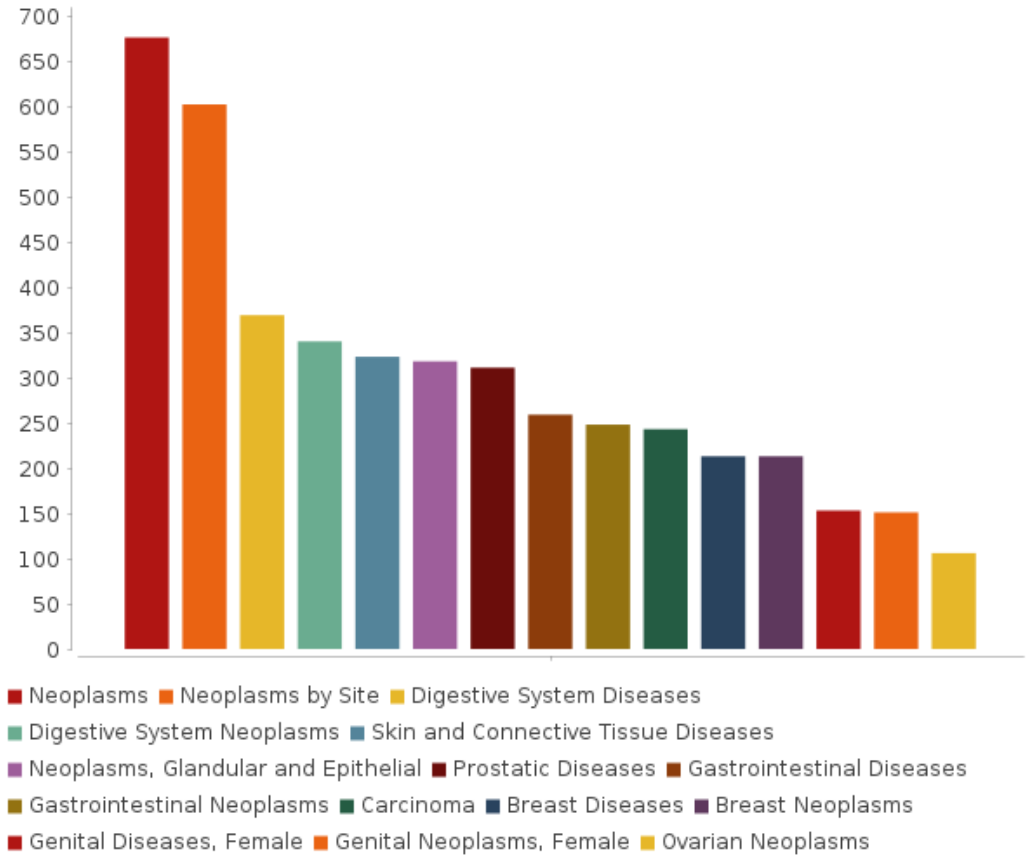
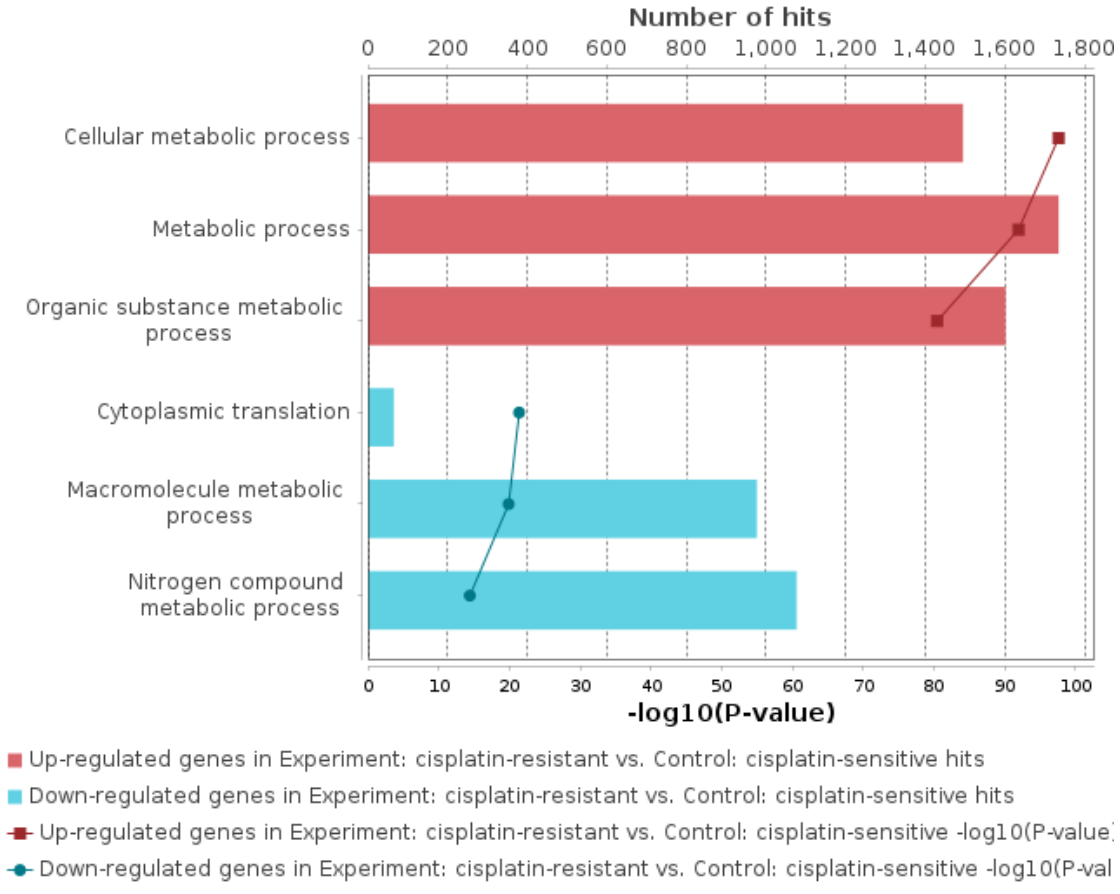


Figure 8. Enriched HumanPSD(TM) disease (2024.2) of down-regulated genes in Experiment: cisplatin-resistant vs. Control: cisplatin-sensitive. The size of the bars correspond to the number of biomarkers of the given disease found among the input set.

Full classification →

The result of overall Gene Ontology (GO) analysis of the differentially expressed genes of the studied pathology can be summarized by the following diagram, revealing the most significant functional categories overrepresented among the observed (differentially expressed genes):



3.3. Analysis of enriched transcription factor binding sites and composite modules

In the next step a search for transcription factors binding sites (TFBS) was performed in the regulatory regions of the ***target genes*** by using the TF binding motif library of the TRANSFAC® database. We searched for so called **composite modules** that act as potential condition-specific **enhancers** of the ***target genes*** in their upstream regulatory regions (-1000 bp upstream of transcription start site (TSS)) and identify transcription factors regulating activity of the genes through such **enhancers**.

Classically, **enhancers** are defined as regions in the genome that increase transcription of one or several genes when inserted in either orientation at various distances upstream or downstream of the gene [8]. Enhancers typically have a length of several hundreds of nucleotides and are bound by multiple transcription factors in a cooperative manner [9].

In the current work we use the Epigenomics data from the track(s) "GSM385747_CpG_NM.fixed.hg38.top300" to predict positions of potential ***enhancers*** regulating the differentially expressed genes revealed by comparative epigenomics analysis. We took genomic regions -550bp upstream and 550bp downstream from the middle point of each interval of the track and check if these regions are located inside the 5kb flanking areas of the differentially expressed genes (or inside the body of the genes). In such cases, these genomic regions are used for the search for potential condition-specific enhancers. In all other cases when the differentially expressed genes did not contain epigenomic peaks in their body or in the 5kb flanking regions we used the upstream regulatory regions of these genes (-1000bp upstream and 100bp downstream of TSS) for the search for condition-specific enhancers.

We applied the Composite Module Analyst (CMA) [8] method to detect such potential enhancers, as targets of multiple TFs bound in a cooperative manner to the regulatory regions of the genes of interest. CMA applies a genetic algorithm to construct a generalized model of the enhancers by specifying combinations of TF motifs (from TRANSFAC®) whose sites are most frequently clustered together in the regulatory regions of the studied genes. CMA identifies the transcription factors that through their cooperation provide a synergistic effect and thus have a great influence on the gene regulation process.

Enhancer model potentially involved in regulation of target genes (up-regulated genes in Experiment: cisplatin-resistant vs. Control: cisplatin-sensitive).

To build the most specific composite modules we choose top 300 significant up-regulated genes as the input of CMA algorithm. The obtained CMA model is then applied to compute CMA score for all up-regulated genes in Experiment: cisplatin-resistant vs. Control: cisplatin-sensitive.

The model consists of 2 module(s). Below, for each module the following information is shown:
- PWMs producing matches,
- number of individual matches for each PWM,
- score of the best match.

Module 1:

V$NFYA_05	V$PPARGRXRA_02	V$NF1C_02	V$GATA1_10	V$PAX5_05	V$E2F8_04
0.92; N=2	0.81; N=3	0.98; N=1	0.85; N=3	0.79; N=2	0.89; N=2

Module width: 119

Module 2:

V$LEF1_Q5_01	V$IRF8_Q6	V$EP300_11	V$MEIS1DLX2_01	V$TIEG1_Q6	V$RBPJK_03
0.96; N=1	0.97; N=1	0.99; N=1	0.87; N=1	0.98; N=1	0.96; N=1

Module width: 145

Model score (-p*log10(pval)): 13.90
Wilcoxon p-value (pval): 5.10e-30
Penalty (p): 0.475
Average yes-set score: 4.71
Average no-set score: 3.70
AUC: 0.74
Separation point: 4.29
False-positive: 30.80%
False-negative: 32.00%

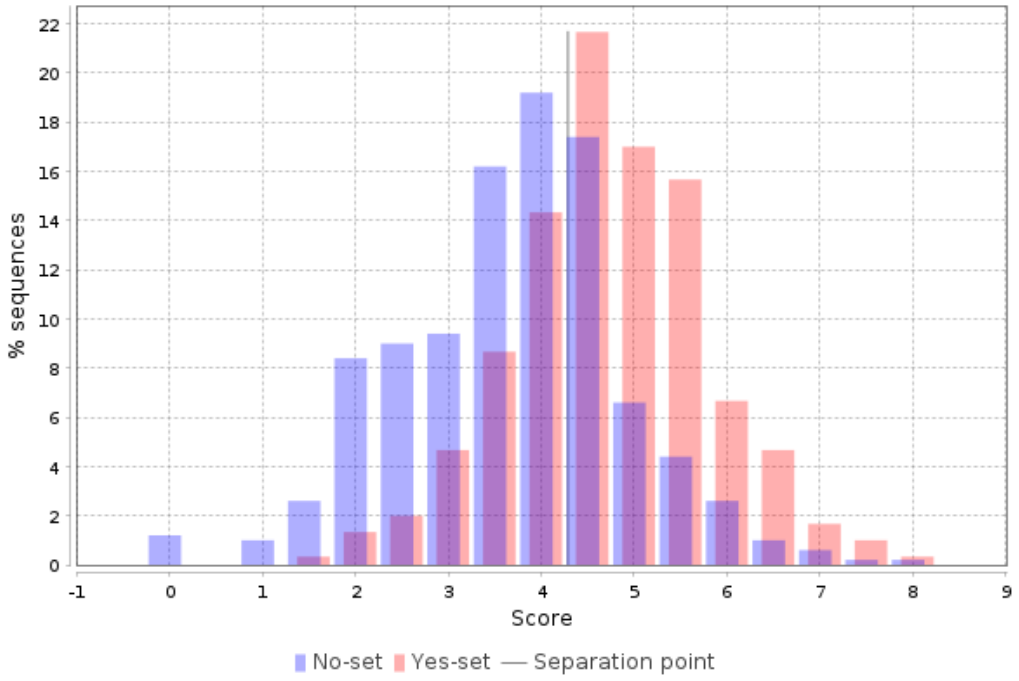


Table 4. List of top ten up-regulated genes in Experiment: cisplatin-resistant vs. Control: cisplatin-sensitive with identified enhancers in their regulatory regions. ***CMA score*** *- the score of the CMA model of the enhancer identified in the regulatory region.*

See full table →

Ensembl IDs	Gene symbol	Gene description	CMA score	Factor names
ENSG00000213719	CLIC1	chloride intracellular channel 1	9.2	E2F-8(h), NF-YA(h), PPARgamma(h),RXRalpha(h), p300(h), GATA-1(h), NF-1C(h), LEF-1(h)...
ENSG00000112640	PPP2R5D	protein phosphatase 2 regulatory subunit B'delta	8.27	PPARgamma(h),RXRalpha(h), GATA-1(h), NF-1C(h), NF-YA(h), p300(h), KLF10(h), LEF-1(h)
ENSG00000215769	ARHGAP27P1-BPTFP1-KPNA2P3	ARHGAP27P1-BPTFP1-KPNA2P3 readthrough, transcribed pseudogene	8.21	p300(h), KLF10(h), E2F-8(h), NF-1C(h), PPARgamma(h),RXRalpha(h), RBP-Jkappa(h), GATA-1(h)
ENSG00000183354	BRD10	bromodomain containing 10	8.15	PPARgamma(h),RXRalpha(h), KLF10(h), E2F-8(h), NF-YA(h), RBP-Jkappa(h), p300(h), IRF-8(h)...
ENSG00000094880	CDC23	cell division cycle 23	7.99	NF-YA(h), E2F-8(h), PPARgamma(h),RXRalpha(h), RBP-Jkappa(h), IRF-8(h), GATA-1(h), LEF-1(h)...
ENSG00000160752	FDPS	farnesyl diphosphate synthase	7.93	PPARgamma(h),RXRalpha(h), LEF-1(h), IRF-8(h), GATA-1(h), RBP-Jkappa(h), NF-YA(h), NF-1C(h)...
ENSG00000160014	CALM3	calmodulin 3	7.91	LEF-1(h), IRF-8(h), p300(h), RBP-Jkappa(h), E2F-8(h), PPARgamma(h),RXRalpha(h), NF-1C(h)...
ENSG00000197043	ANXA6	annexin A6	7.83	DLX-2(h),Meis1(h), IRF-8(h), PPARgamma(h),RXRalpha(h), E2F-8(h), LEF-1(h), NF-1C(h), GATA-1(h)...
ENSG00000169221	TBC1D10B	TBC1 domain family member 10B	7.75	PPARgamma(h),RXRalpha(h), NF-1C(h), NF-YA(h), p300(h), DLX-2(h),Meis1(h), LEF-1(h), KLF10(h)
ENSG00000014641	MDH1	malate dehydrogenase 1	7.73	p300(h), KLF10(h), RBP-Jkappa(h), LEF-1(h), PPARgamma(h),RXRalpha(h), NF-1C(h), GATA-1(h)...

Enhancer model potentially involved in regulation of target genes (down-regulated genes in Experiment: cisplatin-resistant vs. Control: cisplatin-sensitive).

To build the most specific composite modules we choose top 300 significant down-regulated genes as the input of CMA algorithm. The obtained CMA model is then applied to compute CMA score for all down-regulated genes in Experiment: cisplatin-resistant vs. Control: cisplatin-sensitive.

The model consists of 2 module(s). Below, for each module the following information is shown:
- PWMs producing matches,
- number of individual matches for each PWM,
- score of the best match.

Module 1:

V$ZNF462_01	V$MEIS1EOMES_01	V$RXRA_16	V$TALLIKE_Q6	V$TFAP2CHES7_02	V$TEF_Q6
0.95; N=2	0.90; N=2	0.91; N=1	0.97; N=2	0.88; N=2	0.91; N=1

V$EN2_03
0.98; N=2

Module width: 132

Module 2:

V$NFAT4_Q5	V$TCF4_06	V$RSRFC4_Q2	V$RELA_10	V$HSF1_01	V$E2F2_11
1.00; N=3	0.91; N=1	0.88; N=2	0.92; N=2	0.94; N=2	0.87; N=2

Module width: 116

Model score (-p*log10(pval)): 15.30
Wilcoxon p-value (pval): 9.49e-34
Penalty (p): 0.463
Average yes-set score: 3.78
Average no-set score: 2.32
AUC: 0.75
Separation point: 2.80
False-positive: 35.60%
False-negative: 25.00%

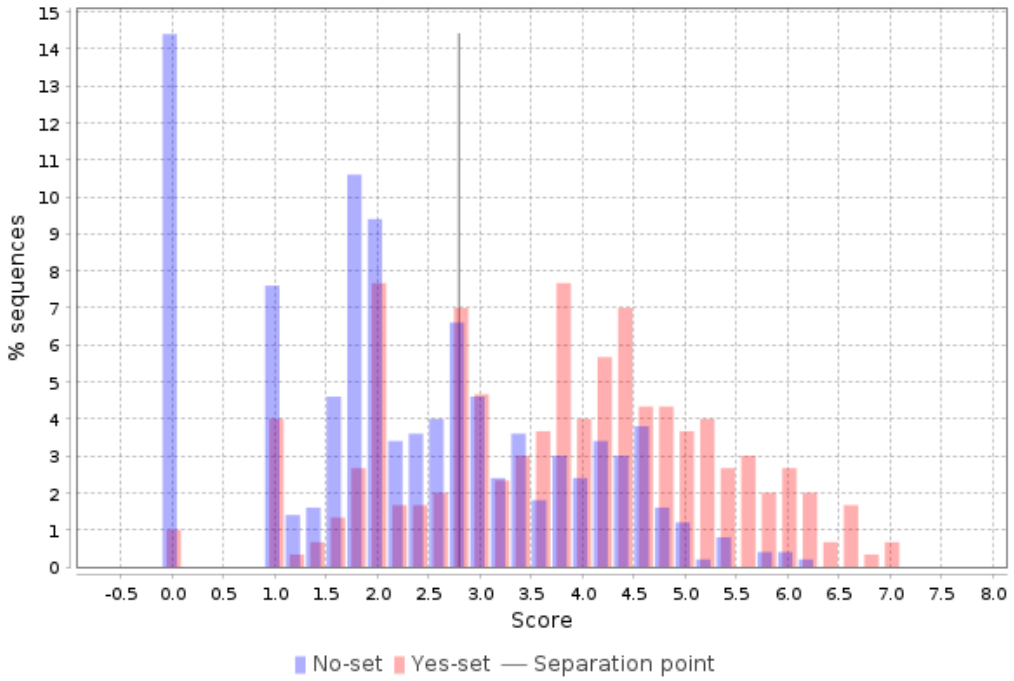


Table 5. List of top ten down-regulated genes in Experiment: cisplatin-resistant vs. Control: cisplatin-sensitive with identified enhancers in their regulatory regions. ***CMA score*** *- the score of the CMA model of the enhancer identified in the regulatory region.*

See full table →

Ensembl IDs	Gene symbol	Gene description	CMA score	Factor names
ENSG00000164056	SPRY1	sprouty RTK signaling antagonist 1	8.12	NFATc3(h), HSF1(h), TCF-7L2(h), EN-2(h), ZNF462(h), RXRalpha(h), TEF(h)
ENSG00000089876	DHX32	DEAH-box helicase 32 (putative)	7.29	E2F-2(h), ZNF462(h), NF-kappaB-p65(h), HSF1(h), EN-2(h), Meis1(h),TBR-2(h)
ENSG00000133056	PIK3C2B	phosphatidylinositol-4-phosphate 3-kinase catalytic subunit type 2 beta	7.21	ZNF462(h), RXRalpha(h), HSF1(h), NFATc3(h), TCF-7L2(h)
ENSG00000215808	LINC01139	long intergenic non-protein coding RNA 1139	7.04	NFATc3(h), RXRalpha(h), HSF1(h), ZNF462(h), NF-kappaB-p65(h), EN-2(h)
ENSG00000101132	PFDN4	prefoldin subunit 4	7	TCF-7L2(h), E2F-2(h), ZNF462(h), NF-kappaB-p65(h), EN-2(h), RXRalpha(h), Meis1(h),TBR-2(h)...
ENSG00000112186	CAP2	cyclase associated actin cytoskeleton regulatory protein 2	6.95	Meis1(h),TBR-2(h), ZNF462(h), EN-2(h), RXRalpha(h), NFATc3(h), TCF-7L2(h), TEF(h)
ENSG00000143162	CREG1	cellular repressor of E1A stimulated genes 1	6.94	NFATc3(h), TCF-7L2(h), ZNF462(h), Meis1(h),TBR-2(h), HEN2(h),Lyl-1(h),Tal-1(h), HSF1(h), RXRalpha(h)...
ENSG00000263465	SRSF8	serine and arginine rich splicing factor 8	6.93	MEF-2A(h), EN-2(h), RXRalpha(h), TEF(h), NFATc3(h), HSF1(h), TCF-7L2(h)
ENSG00000152990	ADGRA3	adhesion G protein-coupled receptor A3	6.93	ZNF462(h), HEN2(h),Lyl-1(h),Tal-1(h), AP-2gamma(h),HES-7(h), E2F-2(h), HSF1(h), RXRalpha(h), NF-kappaB-p65(h)
ENSG00000109265	CRACD	capping protein inhibiting regulator of actin dynamics	6.9	EN-2(h), MEF-2A(h), NFATc3(h), ZNF462(h), NF-kappaB-p65(h), HSF1(h), Meis1(h),TBR-2(h)...

On the basis of the enhancer models we identified transcription factors potentially regulating the ***target genes*** of our interest. We found 14 and 17 transcription factors controlling expression of up- and down-regulated genes respectively (see Tables 6-7).

Table 6. Transcription factors of the predicted enhancer model potentially regulating the differentially expressed genes (up-regulated genes in Experiment: cisplatin-resistant vs. Control: cisplatin-sensitive). ***Yes-No ratio*** *is the ratio between frequencies of the sites in Yes sequences versus No sequences. It describes the level of the enrichment of binding sites for the indicated TF in the regulatory target regions.* ***Regulatory score*** *is the measure of involvement of the given TF in the controlling of expression of genes that encode master regulators presented below (through positive feedback loops).*

See full table →

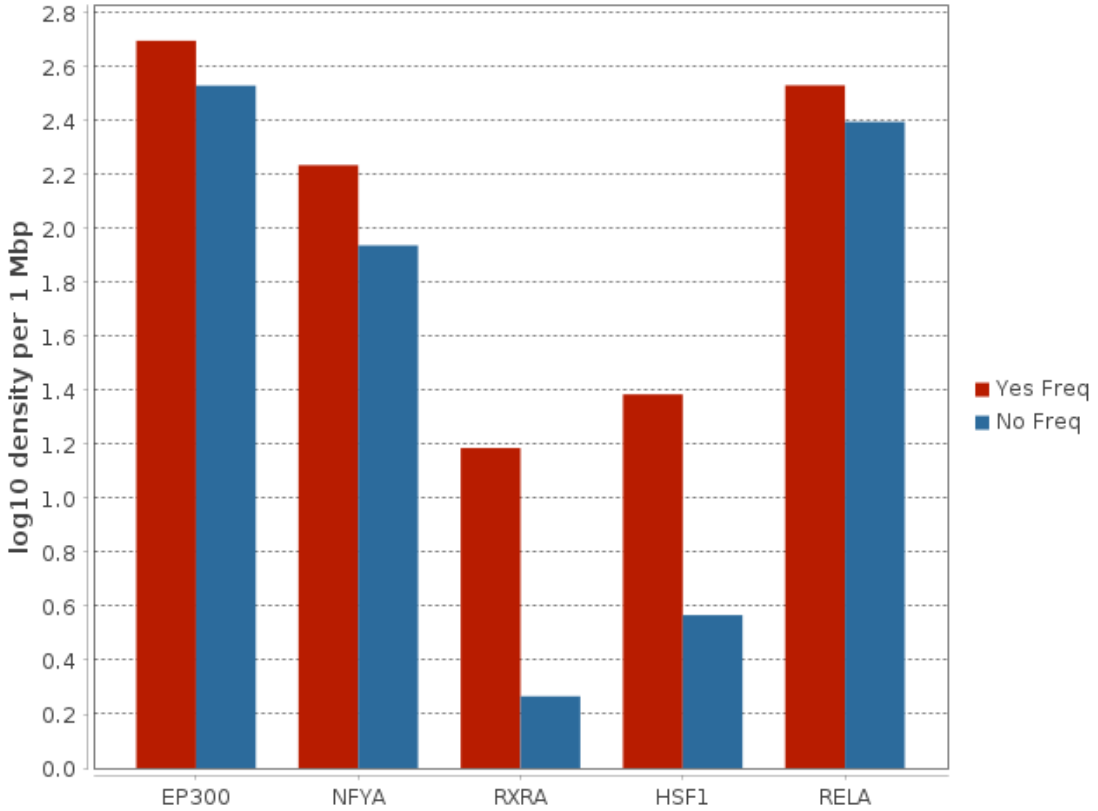
ID	Gene symbol	Gene description	Regulatory score	Yes-No ratio
MO000056654	EP300	E1A binding protein p300	3.73	1.47
MO000025939	NFYA	nuclear transcription factor Y subunit alpha	3.26	1.98
MO000019619	RXRA	retinoid X receptor alpha	3.24	8.33
MO000033565	PPARG	peroxisome proliferator activated receptor gamma	3.12	1.86
MO000159782	LEF1	lymphoid enhancer binding factor 1	2.91	1.89
MO000046001	GATA1	GATA binding protein 1	2.85	1.67
MO000024683	PAX5	paired box 5	2.81	1.53
MO000030964	RBPJ	recombination signal binding protein for immunoglobulin kappa J region	2.54	10.83
MO000023424	IRF8	interferon regulatory factor 8	2.11	1.46
MO000024750	NFIC	nuclear factor I C	1.39	2.68

Table 7. Transcription factors of the predicted enhancer model potentially regulating the differentially expressed genes (down-regulated genes in Experiment: cisplatin-resistant vs. Control: cisplatin-sensitive). ***Yes-No ratio*** *is the ratio between frequencies of the sites in Yes sequences versus No sequences. It describes the level of the enrichment of binding sites for the indicated TF in the regulatory target regions.* ***Regulatory score*** *is the measure of involvement of the given TF in the controlling of expression of genes that encode master regulators presented below (through positive feedback loops).*

See full table →

ID	Gene symbol	Gene description	Regulatory score	Yes-No ratio
MO000033378	HSF1	heat shock transcription factor 1	4.06	6.59
MO000079319	RELA	RELA proto-oncogene, NF-kB subunit	4	1.37
MO000019619	RXRA	retinoid X receptor alpha	3.65	6.59
MO000032489	TAL1	TAL bHLH transcription factor 1, erythroid differentiation factor	3.51	4.94
MO000026882	TCF7L2	transcription factor 7 like 2	3.42	9.88
MO000004278	E2F2	E2F transcription factor 2	2.78	1.54
MO000092587	ZNF462	zinc finger protein 462	2.77	1.25
MO000020739	NFATC3	nuclear factor of activated T cells 3	2.77	2.34
MO000028249	EOMES	eomesodermin	2.57	1.46
MO000025819	LYL1	LYL1 basic helix-loop-helix family member	2.54	1.62

The following diagram represents the key transcription factors, which were predicted to be potentially regulating differentially expressed genes in the analyzed pathology: EP300, NFYA, RXRA, HSF1 and RELA.



3.4. Finding master regulators in networks

In the second step of the upstream analysis common regulators of the revealed TFs were identified. These master regulators appear to be the key candidates for therapeutic targets as they have a master effect on regulation of intracellular pathways that activate the pathological process of our study. The identified master regulators are shown in Tables 8-9.

*Table 8. Master regulators that may govern the regulation of **up-regulated** genes in Experiment: cisplatin-resistant vs. Control: cisplatin-sensitive. **Total rank** is the sum of the ranks of the master molecules sorted by keynode score, CMA score, transcriptomics and epigenomics data.*
See full table →

ID	Master molecule name	Gene symbol	Gene description	logFC	Total rank
MO000010977	PDGFRalpha(h)	PDGFRA	platelet derived growth factor receptor alpha	2.83	80
MO000112248	PDGFRalpha-isoform1(h)	PDGFRA	platelet derived growth factor receptor alpha	2.83	94
MO000256763	PDGFRalpha-isoform2(h)	PDGFRA	platelet derived growth factor receptor alpha	2.83	94
MO000256764	PDGFRalpha-isoform3(h)	PDGFRA	platelet derived growth factor receptor alpha	2.83	94
MO000159408	vrk1(h)	VRK1	VRK serine/threonine kinase 1	0.98	314
MO000092591	Cdk1-isoform1(h):cyclinB1-isoform1(h)	CCNB1, CDK1	cyclin B1, cyclin dependent kinase 1	0.82	349
MO000021736	CDK2(h)	CDK2	cyclin dependent kinase 2	0.8	375
MO000023445	Cdc25A(h)	CDC25A	cell division cycle 25A	0.78	390
MO000041170	EAC(h)	CYLD	CYLD lysine 63 deubiquitinase	0.95	391
MO000129050	EAC-isoform1(h)	CYLD	CYLD lysine 63 deubiquitinase	0.95	415

*Table 9. Master regulators that may govern the regulation of **down-regulated** genes in Experiment: cisplatin-resistant vs. Control: cisplatin-sensitive. **Total rank** is the sum of the ranks of the master molecules sorted by keynode score, CMA score, transcriptomics and epigenomics data.*
See full table →

ID	Master molecule name	Gene symbol	Gene description	logFC	Total rank
MO000033272	SGK-1(h)	SGK1	serum/glucocorticoid regulated kinase 1	-1	211
MO000005412	Fyn(h)	FYN	FYN proto-oncogene, Src family tyrosine kinase	-0.82	263
MO000273747	SGK-1(h){pT256}{pS422}	SGK1	serum/glucocorticoid regulated kinase 1	-1	309
MO000170234	RNF4(h)	RNF4	ring finger protein 4	-0.9	324
MO000022222	MKP-1(h)	DUSP1	dual specificity phosphatase 1	-1.38	351
MO000200584	sharpin-isoform1(h):rbck1-isoform1(h):HOIP-isoform1(h)	RBCK1, RNF31, SHARPIN	RANBP2-type and C3HC4-type zinc finger containing 1, SHANK associated RH domain interactor, ring fin...	-0.81	358
MO001087366	scmh1:BMI1,PCGF2:PHC1,PHC2,PHC3:RING1,RNF2:CBX2,CBX4,CBX8	BMI1, CBX2, CBX4, CBX8, PCGF2, PHC1, PHC2, PHC3, RING1, RNF2, SCMH1	BMI1 proto-oncogene, polycomb ring finger, Scm polycomb group protein homolog 1, chromobox 2, chromo...	-0.82	361
MO000092896	p38beta(h)	MAPK11	mitogen-activated protein kinase 11	-0.41	412
MO000022209	p38beta1(h)	MAPK11	mitogen-activated protein kinase 11	-0.41	448
MO000016677	EGFR(h)	EGFR	epidermal growth factor receptor	-0.85	473

The intracellular regulatory pathways controlled by the above-mentioned master regulators are depicted in Figures 9 and 10. These diagrams display the connections between identified transcription factors, which play important roles in the regulation of differentially expressed genes, and selected master regulators, which are responsible for the regulation of these TFs.

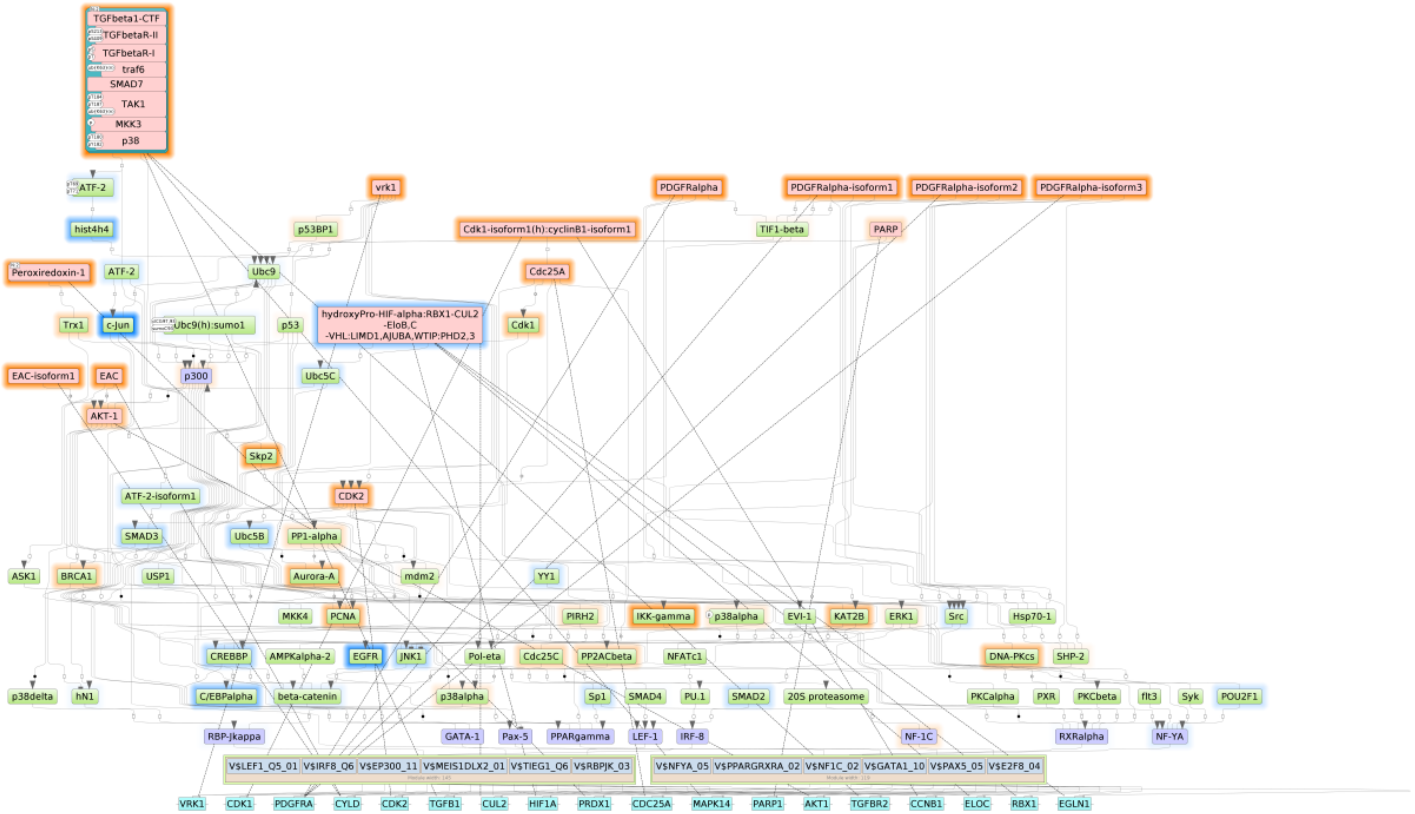


Figure 9. Diagram of intracellular regulatory signal transduction pathways of up-regulated genes in Experiment: cisplatin-resistant vs. Control: cisplatin-sensitive. Master regulators are indicated by red rectangles, transcription factors are blue rectangles, and green rectangles are intermediate molecules, which have been added to the network during the search for master regulators from selected TFs. Orange and blue frames highlight molecules that are encoded by up- and downregulated genes, resp.

See full diagram →

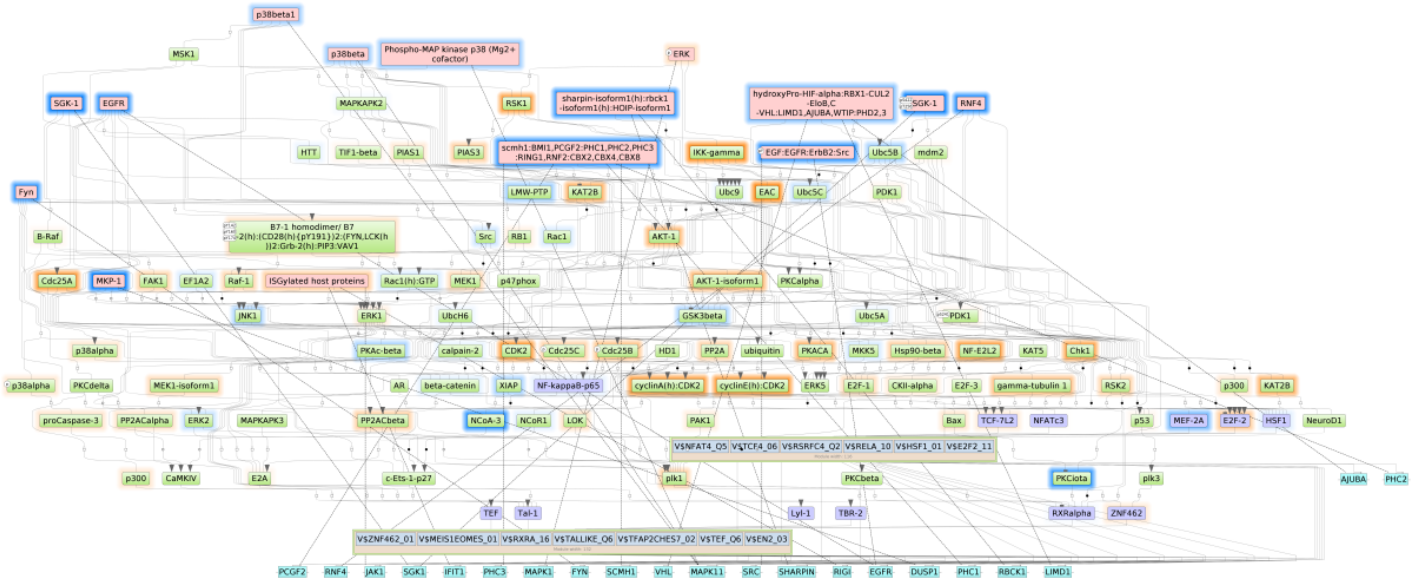


Figure 10. Diagram of intracellular regulatory signal transduction pathways of down-regulated genes in Experiment: cisplatin-resistant vs. Control: cisplatin-sensitive. Master regulators are indicated by red rectangles, transcription factors are blue rectangles, and green rectangles are intermediate molecules, which have been added to the network during the search for master regulators from selected TFs. Orange and blue frames highlight molecules that are encoded by up- and downregulated genes, resp.

See full diagram →

4. Finding prospective drug targets

The identified master regulators that may govern pathology associated genes were checked for druggability potential using HumanPSD™ [5] database of gene-disease-drug assignments and PASS [12-14] software for prediction of biological activities of chemical compounds on the basis of a (Q)SAR approach. Respectively, for each master regulator protein we have computed two Druggability scores: HumanPSD Druggability score and PASS Druggability score. Where Druggability score represents the number of drugs that are potentially suitable for inhibition (or activation) of the corresponding target either according to the information extracted from medical

literature (from HumanPSD™ database) or according to cheminformatics predictions of compounds activity against the examined target (from PASS software).

The cheminformatics druggability check is done using a pre-computed database of spectra of biological activities of chemical compounds from a library of all small molecular drugs from HumanPSD™ database, 2507 pharmaceutically active known chemical compounds in total. The spectra of biological activities has been computed using the program PASS [12-14] on the basis of a (Q)SAR approach.

If both Druggability scores were below defined thresholds (see Methods section for the details) such master regulator proteins were not used in further analysis of drug prediction.

As a result we created the following two tables of prospective drug targets (top targets are shown here):

*Table 10. Prospective drug targets selected from full list of identified master regulators filtered by Druggability score from HumanPSD™ database. **Druggability score** contains the number of drugs that are potentially suitable for inhibition (or activation) of the target. The drug targets are sorted according to the **Total rank** which is the sum of three ranks computed on the basis of the three scores: keynode score, CMA score and expression change score (logFC, if present). See Methods section for details.*

See full table →

Gene symbol	Gene Description	Druggability score	logFC	Total rank
PDGFRA	platelet derived growth factor receptor alpha	55	2.83	80
CCNB1	cyclin B1	18	0.82	349
CDK1	cyclin dependent kinase 1	31	0.82	349
CDK2	cyclin dependent kinase 2	68	0.8	375
CDC25A	cell division cycle 25A	3	0.78	390
AURKA	aurora kinase A	43	0.64	464

*Table 11. Prospective drug targets selected from full list of identified master regulators filtered by Druggability score predicted by PASS software. Here, the **Druggability score** for master regulator proteins is computed as a sum of PASS calculated probabilities to be active as a target for various small molecular compounds. The drug targets are sorted according to the **Total rank** which is the sum of three ranks computed on the basis of the three scores: keynode score, CMA score and expression change score (logFC, if present). See Methods section for details.*

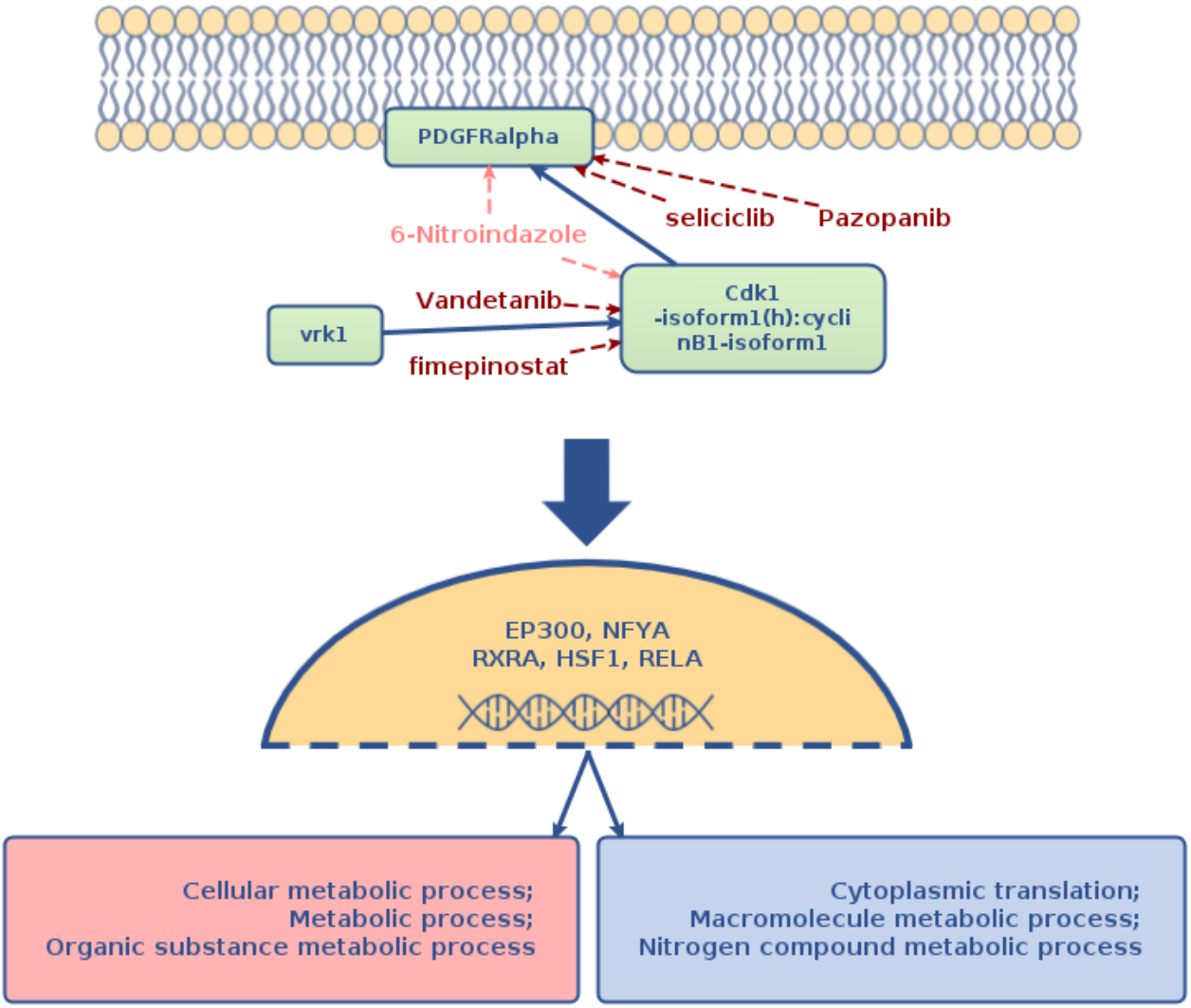
See full table →

Gene symbol	Gene Description	Druggability score	logFC	Total rank
PDGFRA	platelet derived growth factor receptor alpha	6.48	2.83	80
CCNB1	cyclin B1	1.3	0.82	349
CDK1	cyclin dependent kinase 1	5.78	0.82	349
CDK2	cyclin dependent kinase 2	7.76	0.8	375
CDC25A	cell division cycle 25A	3.64	0.78	390
AURKA	aurora kinase A	1.06	0.64	464

Below we represent schematically the main mechanism of the studied pathology. In the schema we considered the top two drug targets of each of the two categories computed above. In addition we have added two top identified master regulators for which no drugs may be identified yet, but that are playing the crucial role in the molecular mechanism of the studied pathology. Thus the molecular mechanism of the studied pathology was predicted to be mainly based on the following key master regulators:

- PDGFRalpha
- vrk1
- Cdk1-isoform1(h):cyclinB1-isoform1

This result allows us to suggest the following schema of affecting the molecular mechanism of the studied pathology:



Drugs which are shown on this schema: Vandetanib, 6-Nitroindazole, seliciclib, fimepinostat and Pazopanib, should be considered as a prospective research initiative for further drug repurposing and drug development. These drugs were selected as top matching treatments to the most prospective drug targets of the studied pathology, however, these results should be considered with special caution and are to be used for research purposes only, as there is not enough clinical information for adapting these results towards immediate treatment of patients.
The drugs given in dark red color on the schema are FDA approved drugs or drugs which have gone through various phases of clinical trials as active treatments against the selected targets.
The drugs given in pink color on the schema are drugs, which were cheminformatically predicted to be active against the selected targets.

5. Identification of potential drugs

In the last step of the analysis we strived to identify known activities as well as drugs with cheminformatically predicted activities that are potentially suitable for inhibition (or activation) of the identified molecular targets in the context of specified human diseases(s).

Proposed drugs are top ranked drug candidates, that were found to be active on the identified targets and were selected from 4 categories:

1. FDA approved drugs or used in clinical trials drugs for the studied pathology;
2. Repurposing drugs used in clinical trials for other pathologies;
3. Drugs, predicted by PASS to be active against identified drug targets and against the studied pathology;
4. Drugs, predicted by PASS to be active against identified drug targets but for other pathologies.

Proposed drugs were selected on the basis of Drug rank which was computed from the ranks sum based on the individual ranks of the following scores:

- Target activity score (depends on ranks of all targets that were found for the selected drug);
- Disease activity score (weighted sum of number of clinical trials on disease(s) under study where the selected drug is known to be applied or PASS Disease activity score - cheminformatically predicted property of the compound to be active against the studied disease(s));
- Clinical validity score (applicable only for drugs predicted on the basis of literature curation in HumanPSD™ database (Tables 13 and 14), reflects the number of the highest clinical trials phase on which the drug was tested for any pathology).

You can refer to the Methods section for more details on drug ranking procedure.

Based on the Drug rank, a numerical value of Drug score was calculated, which reflects the potential activity of the respective drug on the overall molecular mechanism of the studied pathology. Drug score values belong to the range from 1 to 100 and are calculated as a quotient of maximum drug rank and the drug rank of the given drug multiplied by 100.

Top drugs of each category are given in the tables below:

Drugs approved in clinical trials for Oncology

Table 12. Clinically approved (FDA, ENA, etc.) drugs for the studied pathology (most promising and clinically approved treatment candidates selected for the identified drug targets on the basis of literature curation in HumanPSD™ database)
See full table →

Name	Target names	Drug score	Disease activity score	Disease trial phase	Approved
Doxorubicin	MAPK14, NFE2L2, PIK3CB, PIK3CA, TGFB1, TOP2A, BAX, BIRC5, TOP1, BRCA1, CDKN1B	91	5	small molecule,approved,investigational	Ovarian Neoplasms (ClinicalTrials, ClinicalTrials, ClinicalTrials, ClinicalTrials, ClinicalTrials, ClinicalTrials, ClinicalTrials, ClinicalTrials, ClinicalTrials, DailyMed)
Paclitaxel	PIK3CA, TOP2A, CASP3, E2F1, BIRC5, CDK1, CDK2, MAPK3, TUBG1, BRCA1	90	3	small molecule,approved	Ovarian Neoplasms (ClinicalTrials, ClinicalTrials, ClinicalTrials, ClinicalTrials, FDA, FDA)
Gemcitabine	RRM1, ERBB2, HRAS, CHEK1, BRCA1	89	3	small molecule,approved	Ovarian Neoplasms (ClinicalTrials, ClinicalTrials, ClinicalTrials, ClinicalTrials, FDA)
Amifostine	TGFB1	48	2	small molecule,approved,investigational	Ovarian Neoplasms (ClinicalTrials, DailyMed)
Olaparib	PARP1	36	4	small molecule,approved	Ovarian Neoplasms (FDA, FDA)

The ***Disease trial phase*** column reflects the maximum clinical trials phase in which the drug was studied for the analyzed pathology.

Drugs approved in clinical trials

Table 13. Drugs used in clinical trials for the studied pathology (most promising treatment candidates selected for the identified drug targets on the basis of literature curation in HumanPSD™ database)
See full table →

Name	Target names	Drug score	Disease activity score	Disease trial phase
Pazopanib	STK10, RPS6KA3, CAMK2G, MET, IKBKE, MAPK6, PAK2, PRKACA, TTK, MAP3K11, PDGFRA, RAF1, AURKB, CDK7, RPS6KA1, MAPK14, NEK2, PLK1, NEK6, ERBB2, AKT1, AURKA, CHEK1, MAP3K20, MAPK3, MAP2K6, TYK2, MELK, PIM2, CDK8, MAPK9, CDK9, PRKD3, TYRO3, PDGFRB, MAP2K1, MAP4K3, WEE1, PTK2, PIK3CA, TGFBR2, PAK1, CDK5, CSNK2A1, CDK2, PRKD2, PLK4, STK3	96	2	small molecule,approved
Gefitinib	STK10, RPS6KA3, CAMK2G, MET, IKBKE, MAPK6, PAK2, PRKACA, TTK, MAP3K11, PDGFRA, RAF1, AURKB, CDK7, RPS6KA1, MAPK14, NEK2, PLK1, NEK6, ERBB2, AKT1, AURKA, CHEK1, MAP3K20, MAPK3, MAP2K6, TYK2, MELK, PIM2, CDK8, MAPK9, CDK9, PRKD3, TYRO3, PDGFRB, MAP2K1, MAP4K3, WEE1, PTK2, PIK3CA, TGFBR2, PAK1, CDK5, CSNK2A1, CDK2, PRKD2, PLK4, STK3	95	2	small molecule,approved,investigational
Imatinib	STK10, IKBKE, PAK2, PRKACA, MAP3K11, NEK2, NEK6, ERBB2, MAPK3, MAP2K6, TYK2, MELK, PRKD3, PDGFRB, MAP2K1, MAP4K3, WEE1, PTK2, CSNK2A1, PLK4, STK3, RPS6KA3, CAMK2G, MET, MAPK6, TTK, PDGFRA, RAF1, AURKB, CDK7, RPS6KA1, MAPK14, PLK1, AKT1, AURKA, CHEK1, MAP3K20, BIRC5, PIM2, CDK8, MAPK9, CDK9, TYRO3, PIK3CA, TGFBR2, PAK1, CDK5, CDK2, PRKD2	95	2	small molecule,approved
Vandetanib	STK10, IKBKE, PAK2, PRKACA, MAP3K11, NEK2, NEK6, ERBB2, MAPK3, MAP2K6, TYK2, MELK, PRKD3, PDGFRB, MAP2K1, MAP4K3, WEE1, PTK2, CSNK2A1, PLK4, STK3, RPS6KA3, CAMK2G, MET, MAPK6, TTK, CDK1, PDGFRA, RAF1, AURKB, CDK7, RPS6KA1, MAPK14, PLK1, AKT1, AURKA, CHEK1, MAP3K20, PIM2, CDK8, MAPK9, CDK9, TYRO3, PIK3CA, TGFBR2, PAK1, CDK5, CDK2, PRKD2	95	1	small molecule,approved
Erlotinib	STK10, IKBKE, PAK2, PRKACA, MAP3K11, NEK2, NEK6, ERBB2, MAPK3, MAP2K6, TYK2, MELK, PRKD3, PDGFRB, MAP2K1, MAP4K3, WEE1, PTK2, CSNK2A1, PLK4, STK3, RPS6KA3, CAMK2G, MET, MAPK6, TTK, PDGFRA, RAF1, AURKB, CDK7, RPS6KA1, ILK, MAPK14, PLK1, AKT1, AURKA, CHEK1, MAP3K20, BIRC5, PIM2, CDK8, MAPK9, CDK9, TYRO3, PIK3CA, TGFBR2, PAK1, CDK5, CDK2, PRKD2	95	1	small molecule,approved,investigational

The ***Disease trial phase*** column reflects the maximum clinical trials phase in which the drug was studied for the analyzed pathology.

Repurposing drugs

Table 14. Repurposed drugs used in clinical trials for other pathologies (prospective drugs against the identified drug targets on the basis of literature curation in HumanPSD™ database)
See full table →

Name	Target names	Drug score	Maximum trial phase
fimepinostat	HK2, CCNB1, AKT1S1, CDK1, CDC25C, RB1, AURKB, CCND2, CDKN1B, MAPK14, XRCC6, PARP1, PLK1, HSPA4, AKT1, BAX, E2F1, AURKA, CHEK1, BIRC5, MAPK3, BRCA1, MAPK9, FOXM1, PDGFRB, TP53BP1, PDK1, RAD51, WEE1, RRM2, CASP3, CDK2	90	EARLY_PHASE1: Astrocytoma, Diffuse Intrinsic Pontine Glioma, Glioblastoma, Glioma, Medulloblastoma, Recurrence
Curcumin	CAMK2G, MET, MAPK6, HK2, CCNB1, CDK1, CDC25C, CDKN1B, CCNE2, CDC20, MAPK14, NFE2L2, PARP1, ARNT, STAT3, HSPA4, HIF1A, AKT1, BAX, CHEK1, BIRC5, TAFAZZIN, MAPK3, CHUK, ATR, RHOA, BECN1, PCNA, YWHAE, MAPK9, CASP7, FOXM1, TGFB1, DNMT3A, SKP2, IGFBP5, APH1A, DNMT3B, CASP3, PAK1, DNMT1, CDKN1C, CCNA2, CDK2, CHEK2, CLTC, JAG1	89	EARLY_PHASE1: Chronic Periodontitis, Hematoma, Hematoma, Subdural, Hematoma, Subdural, Chronic, Hematoma, Subdural, Intracranial, Oral Ulcer, Periodontal Pocket, Periodontitis, Recurrence, Ulcer
seliciclib	STK10, IKBKE, PAK2, PRKACA, MAP3K11, NEK2, NEK6, ERBB2, MAPK3, MAP2K6, TYK2, MELK, PRKD3, PDGFRB, MAP2K1, MAP4K3, WEE1, PTK2, CSNK2A1, CHEK2, PLK4, STK3, RPS6KA3, CAMK2G, MET, MAPK6, TTK, CDK1, PDGFRA, RAF1, AURKB, CDK7, RPS6KA1, MAPK14, PLK1, AKT1, AURKA, CHEK1, MAP3K20, HIPK2, PIM2, CDK8, MAPK9, CDK9, TYRO3, PIK3CA, TGFBR2, PAK1, CDK5, CDK2, PRKD2	89	PHASE2: Cystic Fibrosis, Cysts, Fibrosis
midostaurin	STK10, IKBKE, PAK2, PRKACA, MAP3K11, NEK2, HSPA4, NEK6, ERBB2, MAPK3, MAP2K6, TYK2, MELK, PRKD3, PDGFRB, MAP2K1, MAP4K3, PIK3CB, WEE1, PTK2, CASP3, CSNK2A1, PLK4, STK3, RPS6KA3, CAMK2G, MET, MAPK6, TTK, PDGFRA, RAF1, AURKB, CDK7, RPS6KA1, MAPK14, PLK1, AKT1, AURKA, CHEK1, MAP3K20, PIM2, CDK8, MAPK9, CASP7, CDK9, TYRO3, PIK3CA, TGFBR2, PAK1, CDK5, CDK2, PRKD2	89	PHASE1: Leukemia, Leukemia, Myeloid, Leukemia, Myeloid, Acute, Myelodysplastic Syndromes, Preleukemia, Syndrome
Vatalanib	STK10, IKBKE, PAK2, PRKACA, MAP3K11, NEK2, NEK6, ERBB2, MAPK3, MAP2K6, TYK2, MELK, PRKD3, PDGFRB, MAP2K1, MAP4K3, WEE1, PTK2, CSNK2A1, PLK4, STK3, RPS6KA3, CAMK2G, MET, MAPK6, TTK, CDK1, PDGFRA, RAF1, AURKB, CDK7, RPS6KA1, MAPK14, PLK1, AKT1, AURKA, CHEK1, MAP3K20, PIM2, CDK8, MAPK9, CDK9, TYRO3, PIK3CA, TGFBR2, PAK1, CDK5, CDK2, PRKD2	89	PHASE1: Carcinoid Tumor, Carcinoma, Carcinoma, Islet Cell, Carcinoma, Medullary, Carcinoma, Merkel Cell, Carcinoma, Neuroendocrine, Carcinoma, Non-Small-Cell Lung, Carcinoma, Renal Cell, Gastrinoma, Gastrointestinal Neoplasms, Glioblastoma, Glucagonoma, Insulinoma, Intestinal Neoplasms, Lung Neoplasms, Malignant Carcinoid Syndrome, Melanoma, Neoplasm Metastasis, Neoplasms, Pheochromocytoma, Recurrence, Skin Neoplasms, Somatostatinoma

The ***Maximum trial phase*** column reflects the maximum clinical trials phase in which the drug was studied for any pathology.

Table 15. Prospective drugs, predicted by PASS software to be active against the identified drug targets with predicted activity against the studied disease(s) (drug candidates predicted with the cheminformatics tool PASS)
See full table →

Name	Target names	Drug score	Target activity score
Paclitaxel	TUBA1C, TUBA3C, TUBA1A, TUBB6, TUBA1B, TUBB3, TUBB2A, TUBG1	98	1.79
Docetaxel	TUBA1C, TUBA3C, TUBA1A, TUBB6, TUBA1B, TUBB3, TUBB2A, TUBG1	98	1.78

Table 16. Prospective drugs, predicted by PASS software to be active against the identified drug targets, though without cheminformatically predicted activity against the studied disease(s) (drug candidates predicted with the cheminformatics tool PASS)

See full table →

Name	Target names	Drug score	Target activity score
Etoposide	TUBA3C, TUBB, TUBB6, PKM, TOP2A, TUBB2A, TUBG1, TOP1, TUBA1C, TUBA1A, HAT1, HIF1A, TUBA1B, TUBB3, CASP3, LIG1	98	4.23
LE-SN38	HIF1A, TOP2A, CASP3, TOP1	97	1.5
Vinblastine	TUBA1C, TUBA3C, TUBB, TUBA1A, TUBB6, HIF1A, TUBA1B, TUBB3, TUBB2A, TUBG1	97	3.56
Vincristine	TUBA1C, TUBA3C, TUBB, TUBA1A, TUBB6, HIF1A, TUBA1B, TUBB3, TUBB2A, TUBG1	97	3.84
Camptothecin	HIF1A, TOP2A, CASP3, TOP1	97	1.5

As the result of drug search we propose the following drugs as most promising candidates for treating the pathology under study: Pazopanib, fimepinostat, Paclitaxel and Etoposide. These drugs were selected for acting on the following targets: PDGFRA, CCNB1, TUBA3C and TOP2A, which were predicted to be active in the molecular mechanism of the studied pathology.

The selected drugs are top ranked drug candidates from each of the four categories of drugs: (1) FDA approved drugs or used in clinical trials drugs for the studied pathology; (2) repurposing drugs used in clinical trials for other pathologies; (3) drugs, predicted by PASS software to be active against the studied pathology; (4) drugs, predicted by PASS software to be repurposed from other pathologies.

Supplementary drug info

In addition to the approved and repurposed drugs proposed by Genome Enhancer, below the ***Supplementary drug info*** table is given, which contains an extended list of drugs used for treatment of neoplasms. Those drugs which were predicted by Genome Enhancer as prospective treatment candidates for the studied case (both approved and repurposed) have a respective ***Predicted Drug Score*** assigned to them. This value on a scale from 1 to 100 reflects the potential activity of the respective drug on the overall molecular mechanism of the studied pathology. The ***Predicted Drug Score*** column contains the "-" (Not Identified) value in case the drug targets of the respective treatment were not found in the molecular mechanism of the studied pathology.

Table 17. Supplementary drug info: extended list of drugs used for treatment of neoplasms with respective drug scores predicted for the studied pathology.

Drug	Disease	Predicted Drug Score
Abarelix	Prostatic Neoplasms	-
Abemaciclib	Breast Neoplasms	60
Abiraterone	Prostatic Neoplasms, Castration-Resistant	-
Abiraterone acetate	Prostatic Neoplasms, Castration-Resistant	-
Acalabrutinib	Lymphoma, Mantle-Cell	-
Acitretin	Psoriasis	18
Ado-trastuzumab emtansine	Breast Neoplasms|Neoplasms	83
Afatinib	Carcinoma, Non-Small-Cell Lung	20
Aflibercept	Colorectal Neoplasms|Diabetic Retinopathy|Edema|Vascular Diseases|Wet Macular Degeneration	1
Alectinib	Carcinoma, Non-Small-Cell Lung	22
Alemtuzumab	Brain Abscess|Leukemia, Lymphocytic, Chronic, B-Cell|Multiple Sclerosis|Multiple Sclerosis, Relapsing-Remitting|Sclerosis	-
Alitretinoin	Sarcoma, Kaposi	-
Alpelisib	Breast Neoplasms	41
Altretamine	Ovarian Neoplasms	-
Aminolevulinic acid	Keratosis|Keratosis, Actinic	-
Anagrelide	Thrombocythemia, Essential|Thrombocytosis	-
Anastrozole	Breast Neoplasms|Hypersensitivity|Obesity|Obesity, Morbid|Recurrence|Weight Loss	-
Apalutamide	Prostatic Neoplasms, Castration-Resistant	-
Aprepitant	Nausea|Neoplasms|Postoperative Nausea and Vomiting	-
Arsenic trioxide	Leukemia, Promyelocytic, Acute	75
Atezolizumab	Carcinoma, Non-Small-Cell Lung|Carcinoma, Transitional Cell|Triple Negative Breast Neoplasms	-
Avelumab	Carcinoma, Merkel Cell|Carcinoma, Renal Cell|Carcinoma, Transitional Cell	-
Axitinib	Carcinoma, Renal Cell	81
Azacitidine	Anemia, Refractory|Anemia, Refractory, with Excess of Blasts|Leukemia, Myelomonocytic, Chronic|Myelodysplastic Syndromes|Preleukemia|Syndrome	72

Belinostat	Lymphoma, T-Cell, Peripheral	19
Bendamustine	Leukemia, Lymphocytic, Chronic, B-Cell|Leukemia, Lymphoid	-
Bevacizumab	Breast Neoplasms|Colonic Neoplasms|Colorectal Neoplasms|Corneal Neovascularization|Diabetic Retinopathy|Dilatation, Pathologic|Edema|Epistaxis|Glaucoma|Hemorrhage|Macular Degeneration|Macular Edema|Neoplasm Metastasis|Neoplasms|Neovascularization, Pathologic|Optic Nerve Diseases|Pterygium|Rectal Neoplasms|Retinal Detachment|Retinal Diseases|Retinal Vein Occlusion|Telangiectasia, Hereditary Hemorrhagic|Telangiectasis|Vitreous Hemorrhage	-
Bexarotene	Lymphoma, T-Cell|Lymphoma, T-Cell, Cutaneous	-
Bicalutamide	Prostatic Neoplasms	7
Binimetinib	Melanoma	42
Blinatumomab	Precursor B-Cell Lymphoblastic Leukemia-Lymphoma	-
Bortezomib	Brain Abscess|Glomerulonephritis|Glomerulonephritis, IGA|Kidney Diseases|Multiple Myeloma|Neoplasms, Plasma Cell|Nephritis|Renal Insufficiency	58
Bosutinib	Leukemia, Myelogenous, Chronic, BCR-ABL Positive	79
Brentuximab vedotin	Hodgkin Disease|Lymphoma|Lymphoma, Large-Cell, Anaplastic|Lymphoma, T-Cell, Peripheral	-
Brigatinib	Carcinoma, Non-Small-Cell Lung	57
Buserelin	Prostatic Neoplasms	-
Cabazitaxel	Prostatic Neoplasms, Castration-Resistant	68
Cabergoline	Drug-Related Side Effects and Adverse Reactions|Pituitary Neoplasms	-
Cabozantinib	Thyroid Neoplasms	20
Capecitabine	Breast Neoplasms|Colonic Neoplasms|Colorectal Neoplasms	-
Carboplatin	Carcinoma, Non-Small-Cell Lung|Lung Neoplasms|Neoplasms|Neuroendocrine Tumors|Ovarian Neoplasms|Retinoblastoma	-
Carfilzomib	Multiple Myeloma	69
Carmustine	Astrocytoma|Glioblastoma|Hodgkin Disease|Medulloblastoma|Multiple Myeloma|Neoplasms	7
Ceritinib	Carcinoma, Non-Small-Cell Lung	78
Cetuximab	Colorectal Neoplasms	-
Cinacalcet	Anemia|Calcinosis|Cardiovascular Diseases|Hyperparathyroidism|Hyperparathyroidism, Secondary|Kidney Diseases|Kidney Failure, Chronic|Neoplasm Metastasis|Neoplasms|Parathyroid Neoplasms|Renal Insufficiency|Vascular Calcification|Vascular Diseases|Vision Disorders	-
Cisplatin	Carcinoma, Squamous Cell|Neoplasms|Uterine Cervical Neoplasms|Carcinoma, Non-Small-Cell Lung|Esophageal Neoplasms|Carcinoma	71
Cladribine	Leukemia, Hairy Cell	71
Clofarabine	Precursor Cell Lymphoblastic Leukemia-Lymphoma	61
Cobimetinib	Melanoma	35
Copanlisib	Lymphoma, Follicular	71
Crizotinib	Carcinoma, Non-Small-Cell Lung	43
Cyproterone acetate	Prostatic Neoplasms	-
Dabrafenib	Melanoma	18
Dacomitinib	Carcinoma, Non-Small-Cell Lung	61
Daratumumab	Multiple Myeloma	-
Dasatinib	Leukemia, Myelogenous, Chronic, BCR-ABL Positive|Leukemia, Myeloid, Chronic-Phase|Precursor Cell Lymphoblastic Leukemia-Lymphoma	93
Decitabine	Anemia, Refractory|Anemia, Refractory, with Excess of Blasts|Leukemia, Myelomonocytic, Chronic|Myelodysplastic Syndromes	71
Degarelix	Cardiovascular Diseases|Prostatic Neoplasms|Vascular Diseases	29
Denosumab	Arthritis, Rheumatoid|Bone Diseases|Bone Diseases, Metabolic|Breast Neoplasms|Hyperparathyroidism|Hyperparathyroidism, Primary|Metabolic Diseases|Neoplasm Metastasis|Neoplasms|Osteoporosis|Osteoporosis, Postmenopausal|Prostatic Neoplasms	-
Dexrazoxane	Breast Neoplasms|Cardiomyopathies	32
Dienogest	Menorrhagia	-
Dinutuximab	Neuroblastoma	-
Docetaxel	Breast Neoplasms|Carcinoma, Non-Small-Cell Lung|Prostatic Neoplasms|Squamous Cell Carcinoma of Head and Neck|Stomach Neoplasms	49
Doxorubicin	Neoplasms|Multiple Myeloma|Carcinoma, Ovarian Epithelial|Ovarian Neoplasms|Leukemia, Lymphoid|Breast Neoplasms|Lymphoma, Follicular|Thyroid Neoplasms|Triple Negative Breast Neoplasms|Glioma	91
Durvalumab	Carcinoma, Non-Small-Cell Lung|Carcinoma, Transitional Cell	-
Dutasteride	Alcoholism|Hyperplasia|Hypertrophy|Neoplasms|Prostatic Hyperplasia	-
Duvelisib	Leukemia, Lymphocytic, Chronic, B-Cell|Lymphoma, Follicular	13
Elotuzumab	Multiple Myeloma	29
Enasidenib	Leukemia, Myeloid, Acute	-
Encorafenib	Colorectal Neoplasms|Melanoma	57
Enfortumab vedotin	Carcinoma, Transitional Cell|Neoplasms	19

Entrectinib	Carcinoma, Non-Small-Cell Lung	-
Enzalutamide	Prostatic Neoplasms|Prostatic Neoplasms, Castration-Resistant	-
Epirubicin	Breast Neoplasms	63
Erdafitinib	Urinary Bladder Neoplasms	85
Eribulin	Breast Neoplasms|Drug-Related Side Effects and Adverse Reactions|Neoplasms	19
Erlotinib	Carcinoma, Non-Small-Cell Lung|Neoplasms|Pancreatic Neoplasms	95
Erlotinib hydrochloride	Carcinoma, Non-Small-Cell Lung|Gastrointestinal Stromal Tumors	-
Estramustine	Prostatic Neoplasms	12
Ethinyl Estradiol	Acne Vulgaris|Neoplasms	16
Everolimus	Angiomyolipoma|Arthrogryposis|Astrocytoma|Breast Neoplasms|Carcinoma, Renal Cell|Cysts|Idiopathic Pulmonary Fibrosis|Kidney Diseases, Cystic|Kidney Failure, Chronic|Lipoma|Neuroendocrine Tumors|Primary Graft Dysfunction|Sclerosis|Tuberous Sclerosis	56
Exemestane	Breast Neoplasms	-
Fedratinib	Primary Myelofibrosis	-
Finasteride	Hyperplasia|Neoplasms|Prostatic Hyperplasia	8
Flavopiridol	Leukemia, Lymphocytic, Chronic, B-Cell	89
Fluorouracil	Skin Neoplasms|Neoplasms, Basal Cell|Neoplasms, Second Primary|Neoplasms, Squamous Cell|Neoplasms|Colorectal Neoplasms|Pancreatic Neoplasms	73
Fluoxymesterone	Breast Neoplasms|Hypogonadism|Puberty, Delayed	-
Flutamide	Premenstrual Dysphoric Disorder|Premenstrual Syndrome|Prostatic Neoplasms	42
Fulvestrant	Breast Neoplasms	-
Gefitinib	Carcinoma, Non-Small-Cell Lung	95
Gemcitabine	Breast Neoplasms|Carcinoma, Non-Small-Cell Lung|Ovarian Neoplasms|Pancreatic Neoplasms	89
Gemtuzumab ozogamicin	Leukemia, Myeloid, Acute	-
Gilteritinib	Leukemia, Myeloid, Acute	67
Glasdegib	Leukemia, Myeloid, Acute	-
Goserelin	Atrophy|Breast Neoplasms|Bulbo-Spinal Atrophy, X-Linked|Endometriosis|Muscular Atrophy|Myoma|Prostatic Neoplasms	-
Histrelin	Puberty, Precocious	-
Homoharringtonine	Leukemia, Myelogenous, Chronic, BCR-ABL Positive	85
Ibritumomab	Lymphoma, B-Cell|Lymphoma, Follicular	-
Ibrutinib	Graft vs Host Disease|Leukemia, Lymphocytic, Chronic, B-Cell|Lymphoma, B-Cell, Marginal Zone|Lymphoma, Mantle-Cell|Waldenstrom Macroglobulinemia	68
Idarubicin	Leukemia, Myeloid, Acute	38
Idelalisib	Leukemia, Lymphocytic, Chronic, B-Cell|Lymphoma, Follicular	53
Ifosfamide	Neoplasms	57
Imatinib	Leukemia, Myelogenous, Chronic, BCR-ABL Positive|Mastocytosis, Systemic|Neoplasms	95
Inotuzumab ozogamicin	Precursor B-Cell Lymphoblastic Leukemia-Lymphoma	-
Ipilimumab	Carcinoma, Renal Cell|Melanoma	-
Irinotecan	Colorectal Neoplasms	63
Ivosidenib	Leukemia, Myeloid, Acute	-
Ixabepilone	Breast Neoplasms	43
Ixazomib	Multiple Myeloma	-
Lapatinib	Breast Neoplasms	89
Larotrectinib	Neoplasm Metastasis	46
Lenalidomide	Brain Abscess|Lupus Erythematosus, Cutaneous|Myelodysplastic Syndromes|Neoplasms, Plasma Cell	-
Lenvatinib	Carcinoma, Hepatocellular|Carcinoma, Renal Cell|Thyroid Neoplasms	70
Letrozole	Breast Neoplasms|Cysts|Fibroma|Myofibroma|Myoma|Ovarian Cysts|Syndrome	-
Leuprolide	Hot Flashes|Ovarian Hyperstimulation Syndrome|Prostatic Neoplasms|Puberty, Precocious	-
Levamisole	Ascariasis|Colonic Neoplasms|Helminthiasis	-
Levonorgestrel	Epilepsy|Hyperplasia|Menorrhagia	-
Lomustine	Brain Neoplasms|Hodgkin Disease	-
Lonafarnib	Breast Neoplasms|Carcinoma, Non-Small-Cell Lung|Central Nervous System Neoplasms|Colorectal Neoplasms|Gliosarcoma|Head and Neck Neoplasms|Leukemia, Myelomonocytic, Chronic| Liver Neoplasms|Lymphoma|Myelodysplastic Syndromes|Ovarian Neoplasms|Urethral Neoplasms|Urinary Bladder Neoplasms	23
Lorlatinib	Carcinoma, Non-Small-Cell Lung	10
Masoprocol	Keratosis, Actinic	-
Medroxyprogesterone Acetate	Depression|Depression, Postpartum|Depressive Disorder|Metrorrhagia|Neoplasms|Uterine Hemorrhage	11
Megestrol acetate	Acquired Immunodeficiency Syndrome|Bites and Stings|Breast Neoplasms|Pain|Wasting Syndrome	-

Methotrexate	Neoplasms|Breast Neoplasms|Head and Neck Neoplasms|Ovarian Neoplasms|Lymphoma, T-Cell, Peripheral|Brain Neoplasms|Colorectal Neoplasms|Neuroblastoma|Carcinoma, Squamous Cell	39
Methyltestosterone	Breast Neoplasms|Hypogonadism|Puberty, Delayed	-
Midostaurin	Leukemia, Mast-Cell|Leukemia, Myeloid, Acute|Mastocytosis, Systemic	89
Mitotane	Adrenocortical Carcinoma	-
Mitoxantrone	Autoimmune Diseases|Autoimmune Diseases of the Nervous System|Demyelinating Autoimmune Diseases, CNS|Immune System Diseases|Leukemia, Myeloid, Acute|Multiple Sclerosis|Myelitis|Myelitis, Transverse|Nervous System Diseases|Neuromyelitis Optica|Prostatic Neoplasms, Castration-Resistant	70
Mogamulizumab	Mycosis Fungoides|Neoplasms|Sezary Syndrome	-
Moxetumomab pasudotox	Leukemia, Hairy Cell|Neoplasms	-
Necitumumab	Carcinoma, Non-Small-Cell Lung|Neoplasms	-
Nelarabine	Precursor T-Cell Lymphoblastic Leukemia-Lymphoma	-
Neratinib	Breast Neoplasms	68
Nilotinib	Blast Crisis|Leukemia, Myelogenous, Chronic, BCR-ABL Positive|Leukemia, Myeloid, Chronic-Phase	75
Nilutamide	Prostatic Neoplasms	-
Nintedanib	Fibrosis|Idiopathic Pulmonary Fibrosis	79
Niraparib	Carcinoma, Ovarian Epithelial|Fallopian Tube Neoplasms|Peritoneal Neoplasms	52
Nivolumab	Carcinoma, Non-Small-Cell Lung|Kidney Neoplasms|Neoplasms|Lung Neoplasms|Melanoma	-
Obinutuzumab	Leukemia, Lymphocytic, Chronic, B-Cell	-
Octreotide	Acromegaly|Adenoma|Ascites|Carcinoid Tumor|Fistula|Pancreatic Fistula|Pituitary Diseases|Renal Insufficiency|Vipoma	8
Ofatumumab	Leukemia, Lymphocytic, Chronic, B-Cell	-
Olaparib	Breast Neoplasms|Carcinoma, Ovarian Epithelial|Fallopian Tube Neoplasms|Ovarian Neoplasms|Pancreatic Neoplasms|Peritoneal Neoplasms|Prostatic Neoplasms, Castration-Resistant	36
Olaratumab	Sarcoma	-
Osimertinib	Carcinoma, Non-Small-Cell Lung	64
Oxaliplatin	Colonic Neoplasms|Colorectal Neoplasms|Neoplasms|Rectal Neoplasms	29
Paclitaxel	Acute Coronary Syndrome|Angina Pectoris|Arteriosclerosis|Breast Neoplasms|Carcinoma, Non-Small-Cell Lung|Cardiovascular Diseases|Coronary Artery Disease|Coronary Disease|Coronary Stenosis|Heart Diseases|Myocardial Ischemia|Ovarian Neoplasms|Vascular Diseases	90
Palbociclib	Breast Neoplasms	-
Panitumumab	Colorectal Neoplasms	-
Panobinostat	Multiple Myeloma	1
Pazopanib	Carcinoma|Carcinoma, Renal Cell|Sarcoma	96
Pembrolizumab	Carcinoma, Hepatocellular|Carcinoma, Merkel Cell|Carcinoma, Non-Small-Cell Lung|Carcinoma, Renal Cell|Carcinoma, Transitional Cell|Hodgkin Disease|Melanoma|Neoplasms|Stomach Neoplasms	-
Pemetrexed	Carcinoma, Non-Small-Cell Lung|Mesothelioma	-
Pentostatin	Leukemia, Hairy Cell	-
Pertuzumab	Breast Neoplasms	75
Pomalidomide	Multiple Myeloma	-
Ponatinib	Leukemia, Myelogenous, Chronic, BCR-ABL Positive|Precursor Cell Lymphoblastic Leukemia-Lymphoma	34
Pralatrexate	Lymphoma, T-Cell, Peripheral	-
Radium Ra 223 Dichloride	Prostatic Neoplasms, Castration-Resistant	-
Ramucirumab	Stomach Neoplasms	-
Rasburicase	Hyperuricemia|Leukemia|Lymphoma|Neoplasms|Syndrome|Tumor Lysis Syndrome	-
Regorafenib	Colorectal Neoplasms	80
Relugolix	Prostatic Neoplasms	-
Ribociclib	Breast Neoplasms	-
Rituximab	Arthritis|Arthritis, Rheumatoid|Granulomatosis with Polyangiitis|Leukemia|Leukemia, Lymphoid|Lymphoma|Lymphoma, B-Cell|Lymphoma, Follicular|Lymphoma, Non-Hodgkin|Myelitis|Neuromyelitis Optica|Purpura|Purpura, Thrombocytopenic|Purpura, Thrombocytopenic, Idiopathic|Thrombocytopenia	-
Romidepsin	Lymphoma, T-Cell, Cutaneous	26
Rucaparib	Carcinoma, Ovarian Epithelial|Fallopian Tube Neoplasms|Peritoneal Neoplasms|Prostatic Neoplasms, Castration-Resistant	80
Ruxolitinib	Graft vs Host Disease|Polycythemia|Polycythemia Vera|Primary Myelofibrosis|Thrombocytosis	26
Selinexor	Multiple Myeloma	25
Selumetinib	Neurofibromatosis 1	22
Siltuximab	Giant Lymph Node Hyperplasia	-

Sirolimus	Angiomyolipoma|Constriction, Pathologic|Coronary Restenosis|Eye Diseases|Immune System Diseases|Kidney Failure, Chronic|Lipoma|Tuberous Sclerosis	94
Sonidegib	Carcinoma, Basal Cell	-
Sorafenib	Carcinoma, Hepatocellular|Carcinoma, Renal Cell|Thyroid Neoplasms	94
Sunitinib	Adenoma|Carcinoma, Renal Cell|Digestive System Neoplasms|Gastrointestinal Neoplasms|Gastrointestinal Stromal Tumors|Intestinal Neoplasms	94
Talazoparib	Breast Neoplasms	23
Tamoxifen	Breast Diseases|Cystic Fibrosis|Cysts|Fibroadenoma|Fibrocystic Breast Disease|Hemorrhage|Menorrhagia|Menstruation Disturbances|Metrorrhagia|Neoplasms	59
Tamsulosin	Calculi|Coronary Artery Disease|Heart Diseases|Hernia|Hernia, Inguinal|Inflammation|Ischemia|Lithiasis|Lower Urinary Tract Symptoms|Myocardial Ischemia|Prostatic Hyperplasia|Ureteral Calculi|Urinary Calculi|Urolithiasis|Urologic Diseases	-
Temozolomide	Astrocytoma|Nervous System Neoplasms	-
Temsirolimus	Carcinoma, Renal Cell	74
Teniposide	Precursor Cell Lymphoblastic Leukemia-Lymphoma	70
Thalidomide	Brain Abscess|Immune System Diseases|Multiple Myeloma|Neoplasms, Plasma Cell	-
Tivozanib	Carcinoma, Renal Cell	82
Tocilizumab	Arthritis|Arthritis, Juvenile|Arthritis, Rheumatoid|Behavior|Cytokine Release Syndrome|Giant Cell Arteritis|Neurobehavioral Manifestations|Oral Manifestations|Psychotic Disorders|Schizophrenia|Tic Disorders	-
Topotecan	Small Cell Lung Carcinoma	54
Toremifene	Breast Neoplasms	48
Trabectedin	Leiomyosarcoma|Liposarcoma	-
Trametinib	Carcinoma, Non-Small-Cell Lung|Melanoma	80
Trastuzumab	Breast Neoplasms|Neoplasms	48
Tretinoin	Lentigo	74
Triptorelin	Fatty Liver|Hypogonadism|Infertility, Female|Prostatic Neoplasms	64
Tucatinib	Breast Neoplasms	74
Valrubicin	Urinary Bladder Neoplasms	50
Vandetanib	Thyroid Neoplasms	95
Vemurafenib	Melanoma	54
Venetoclax	Leukemia, Lymphocytic, Chronic, B-Cell|Leukemia, Myeloid, Acute	-
Vinblastine	Glioma	43
Vincristine	Precursor Cell Lymphoblastic Leukemia-Lymphoma	40
Vinorelbine	Carcinoma, Non-Small-Cell Lung	78
Vismodegib	Carcinoma, Basal Cell	-
Vorinostat	Lymphoma, T-Cell, Cutaneous	37
Zoledronate	Arthritis|Bone Marrow Diseases|Brain Abscess|Chronic Kidney Disease-Mineral and Bone Disorder|Chronic Periodontitis|HIV Infections|Hypersensitivity|Infections|Kidney Diseases|Metabolic Diseases|Multiple Myeloma|Neoplasms|Neoplasms, Plasma Cell|Neoplasms, Second Primary|Osteitis|Osteoarthritis|Periodontitis|Pleural Effusion, Malignant|Prostatic Neoplasms|Renal Insufficiency, Chronic|Thalassemia|Wounds and Injuries	-

6. Conclusion

We applied the software package "Genome Enhancer" to a multi-omics data set that contains *transcriptomics and epigenomics* data. The study is done in the context of *Ovarian Neoplasms*. The data were pre-processed, statistically analyzed and differentially expressed genes were identified. Also checked was the enrichment of GO or disease categories among the studied gene sets.

We propose the following drugs as most promising candidates for treating the pathology under study:

These drugs were selected for acting on the following targets: PDGFRA, CCNB1, TUBA3C and TOP2A, which were predicted to be involved in the molecular mechanism of the pathology under study.

The identified molecular mechanism of the studied pathology was predicted to be mainly based on the following key drug targets:

PDGFRalpha, vrk1 and Cdk1-isoform1(h):cyclinB1-isoform1

These potential drug targets should be considered as a prospective research initiative for further drug repurposing and drug development purposes. The following drugs were predicted as, matching those drug targets: Vandetanib, 6-Nitroindazole, seliciclib, fimepinostat and Pazopanib. These drugs should be considered with special caution for research purposes only.

In this study, we came up with a detailed signal transduction network regulating differentially expressed genes in the studied pathology. In this network we have revealed the following top master regulators (signaling proteins and their complexes) that play a crucial role in the molecular mechanism of the studied pathology, which can be proposed as the most promising molecular targets for further drug repurposing and drug development initiatives.

- PDGFRalpha
- vrk1
- Cdk1-isoform1(h):cyclinB1-isoform1

Potential drug compounds which can be affecting these targets can be found in the "Finding prospective drug targets" section.

7. Methods

Databases used in the study

Transcription factor binding sites in promoters and enhancers of differentially expressed genes were analyzed using known DNA-binding motifs described in the TRANSFAC® library, release 2024.2 (geneXplain GmbH, Wolfenbüttel, Germany) (https://genexplain.com/transfac).
The master regulator search uses the TRANSPATH® database (BIOBASE), release 2024.2 (geneXplain GmbH, Wolfenbüttel, Germany) (https://genexplain.com/transpath). A comprehensive signal transduction network of human cells is built by the software on the basis of reactions annotated in TRANSPATH®.
The information about drugs corresponding to identified drug targets and clinical trials references were extracted from HumanPSD™ database, release 2024.2 (https://genexplain.com/humanpsd).
The Ensembl database release Human112.38 (hg38) (http://www.ensembl.org) was used for gene IDs representation and Gene Ontology (GO) (http://geneontology.org) was used for functional classification of the studied gene set.

Methods for the analysis of enriched transcription factor binding sites and composite modules

Transcription factor binding sites in promoters and enhancers of differentially expressed genes were analyzed using known DNA-binding motifs. The motifs are specified using position weight matrices (PWMs) that give weights to each nucleotide in each position of the DNA binding motif for a transcription factor or a group of them.
We search for transcription factor binding sites (TFBS) that are enriched in the promoters and enhancers under study as compared to a background sequence set such as promoters of genes that were not differentially regulated under the condition of the experiment. We denote study and background sets briefly as Yes and No sets. In the current work we used a workflow considering promoter sequences of a standard length of 1100 bp (-1000 to +100). The error rate in this part of the pipeline is controlled by estimating the adjusted p-value (using the Benjamini-Hochberg procedure) in comparison to the TFBS frequency found in randomly selected regions of the human genome (adj.p-value < 0.01).
We have applied the CMA algorithm (Composite Module Analyst) for searching composite modules [7] in the promoters and enhancers of the Yes and No sets. We searched for a composite module consisting of a cluster of 10 TFs in a sliding window of 200-300 bp that statistically significantly separates sequences in the Yes and No sets (minimizing Wilcoxon p-value).

Methods for finding master regulators in networks

We searched for master regulator molecules in signal transduction pathways upstream of the identified transcription factors. The master regulator search uses a comprehensive signal transduction network of human cells. The main algorithm of the master regulator search has been described earlier [3,4]. The goal of the algorithm is to find nodes in the global signal transduction network that may potentially regulate the activity of a set of transcription factors found at the previous step of the analysis. Such nodes are considered as most promising drug targets, since any influence on such a node may switch the transcriptional programs of hundreds of genes that are regulated by the respective TFs. In our analysis, we have run the algorithm with a maximum radius of 12 steps upstream of each TF in the input set. The error rate of this algorithm is controlled by applying it 10000 times to randomly generated sets of input transcription factors of the same set-size. Z-score and FDR value of ranks are calculated then for each potential master regulator node on the basis of such random runs (see detailed description in [9]). We control the error rate by the FDR threshold 0.05.

Methods for analysis of pharmaceutical compounds

We seek for the optimal combination of molecular targets (key elements of the regulatory network of the cell) that potentially interact with pharmaceutical compounds from a library of known drugs and biologically active chemical compounds, using information about known drugs from HumanPSD™ and predicting potential drugs using PASS program.

Method for analysis of known pharmaceutical compounds

We selected compounds from HumanPSD™ database that have at least one target. Next, we sort compounds using "*Drug rank*" that is the sum of the following ranks:

1. ranking by "Target activity score" (*T-score*$_{PSD}$),
2. ranking by "Disease activity score" (*D-score*$_{PSD}$),
3. ranking by "Clinical validity score".

"Target activity score" (*T-score*$_{PSD}$) is calculated as follows:

$$T\text{-}score_{PSD} = -\frac{|T|}{|T| + w(|AT| - |T|))} \sum_{t \in T} log_{10}\left(\frac{rank(t)}{1 + maxRank(T)}\right),$$

where *T* is set of all targets related to the compound intersected with input list, |*T*| is number of elements in *T*, *AT* and |*AT*| are set set of all targets related to the compound and number of elements in it, *w* is weight multiplier, *rank(t)* is rank of given target, *maxRank(T)* equals *max(rank(t))* for all targets *t* in *T*.

We use following formula to calculate "Disease activity score" (*D-score*$_{PSD}$):

$$D\text{-}score_{PSD} = \begin{cases} \sum_{d \in D} \sum_{p \in P} phase(d,p) \\ 0, \ D = \varnothing \end{cases},$$

where *D* is the set of selected diseases, and if *D* is empty set, *D-score*$_{PSD}$*=0*. *P* is a set of all known phases for each disease, *phase(p,d)* equals to the phase number if there are known clinical trials for the selected disease on this phase and zero otherwise.

The clinical validity score reflects the number of the highest clinical trials phase (from 1 to 4) on which the drug was ever tested for any pathology.

Method for prediction of pharmaceutical compounds

In this study, the focus was put on compounds with high pharmacological efficiency and low toxicity. For this purpose, comprehensive library of chemical compounds and drugs was subjected to a SAR/QSAR analysis. This library contains 13040 compounds along with their pre-calculated potential pharmacological activities of those substances, their possible side and toxic effects, as well as the possible mechanisms of action. All biological activities are expressed as probability values for a substance to exert this activity (*Pa*).

We selected compounds that satisfied the following conditions:

1. Toxicity below a chosen toxicity threshold (defines as *Pa*, probability to be active as toxic substance).
2. For all predicted pharmacological effects that correspond to a set of user selected disease(s) *Pa* is greater than a chosen effect threshold.
3. There are at least 2 targets (corresponding to the predicted activity-mechanisms) with predicted *Pa* greater than a chosen target threshold.

The maximum *Pa* value for all toxicities corresponding to the given compound is selected as the "Toxicity score". The maximum *Pa* value for all activities corresponding to the selected diseases for the given compound is used as the "Disease activity score". "Target activity score" (T-score) is calculated as follows:

$$T\text{-}score(s) = \frac{|T|}{|T| + w(|AT| - |T|))} \sum_{m \in M(s)} \left(pa(m) \sum_{g \in G(m)} IAP(g) optWeight(g) \right),$$

where *M(s)* is the set of activity-mechanisms for the given structure (which passed the chosen threshold for activity-mechanisms *Pa*); *G(m)* is the set of targets (converted to genes) that corresponds to the given activity-mechanism (m) for the given compound; *pa(m)* is the probability to be active of the activity-mechanism (m), *IAP(g)* is the invariant accuracy of prediction for gene from *G(m)*; *optWeight(g)* is the additional weight multiplier for gene. *T* is set of all targets related to the compound intersected with input list, |*T*| is number of elements in *T*, *AT* and |*AT*| are set set of all targets related to the compound and number of elements in it, *w* is weight multiplier.

"Druggability score" (D-score) is calculated as follows:

$$D\text{-}score(g) = IAP(g) \sum_{s \in S(g)} \sum_{m \in M(s,g)} pa(m),$$

where *S(g)* is the set of structures for which target list contains given target, *M(s,g)* is the set of activity-mechanisms (for the given structure) that corresponds to the given gene, *pa(m)* is the probability to be active of the activity-mechanism (m), *IAP(g)* is the invariant accuracy of prediction for the given gene.

8. References

1. Kel A, Voss N, Jauregui R, Kel-Margoulis O, Wingender E. Beyond microarrays: Finding key transcription factors controlling signal transduction pathways. *BMC Bioinformatics.* **2006**;7(S2), S13. doi:10.1186/1471-2105-7-s2-s13
2. Stegmaier P, Voss N, Meier T, Kel A, Wingender E, Borlak J. Advanced Computational Biology Methods Identify Molecular Switches for Malignancy in an EGF Mouse Model of Liver Cancer. *PLoS ONE.* **2011**;6(3):e17738. doi:10.1371/journal.pone.0017738
3. Koschmann J, Bhar A, Stegmaier P, Kel A, Wingender E. "Upstream Analysis": An Integrated Promoter-Pathway Analysis Approach to Causal Interpretation of Microarray Data. *Microarrays.* **2015**;4(2):270-286. doi:10.3390/microarrays4020270.
4. Kel A, Stegmaier P, Valeev T, Koschmann J, Poroikov V, Kel-Margoulis OV, and Wingender E. Multi-omics "upstream analysis" of regulatory genomic regions helps identifying targets against methotrexate resistance of colon cancer. *EuPA Open Proteom.* **2016**;13:1-13. doi:10.1016/j.euprot.2016.09.002
5. Michael H, Hogan J, Kel A et al. Building a knowledge base for systems pathology. *Brief Bioinformatics.* **2008**;9(6):518-531. doi:10.1093/bib/bbn038
6. Matys V, Kel-Margoulis OV, Fricke E, Liebich I, Land S, Barre-Dirrie A, Reuter I, Chekmenev D, Krull M, Hornischer K, Voss N, Stegmaier P, Lewicki-Potapov B, Saxel H, Kel AE, Wingender E. TRANSFAC and its module TRANSCompel: transcriptional gene regulation in eukaryotes. *Nucleic Acids Res.* **2006**;34(90001):D108-D110. doi:10.1093/nar/gkj143
7. Kel AE, Gössling E, Reuter I, Cheremushkin E, Kel-Margoulis OV, Wingender E. MATCH: A tool for searching transcription factor binding sites in DNA sequences. *Nucleic Acids Res.* **2003**;31(13):3576-3579. doi:10.1093/nar/gkg585
8. Waleev T, Shtokalo D, Konovalova T, Voss N, Cheremushkin E, Stegmaier P, Kel-Margoulis O, Wingender E, Kel A. Composite Module Analyst: identification of transcription factor binding site combinations using genetic algorithm. *Nucleic Acids Res.* **2006**;34(Web Server issue):W541-5.
9. Krull M, Pistor S, Voss N, Kel A, Reuter I, Kronenberg D, Michael H, Schwarzer K, Potapov A, Choi C, Kel-Margoulis O, Wingender E. TRANSPATH: an information resource for storing and visualizing signaling pathways and their pathological aberrations. *Nucleic Acids Res.* **2006**;34(90001):D546-D551. doi:10.1093/nar/gkj107
10. Boyarskikh U, Pintus S, Mandrik N, Stelmashenko D, Kiselev I, Evshin I, Sharipov R, Stegmaier P, Kolpakov F, Filipenko M, Kel A. Computational master-regulator search reveals mTOR and PI3K pathways responsible for low sensitivity of NCI-H292 and A427 lung cancer cell lines to cytotoxic action of p53 activator Nutlin-3. *BMC Med Genomics.* **2018**;11(1):12. doi:10.1186/1471-2105-7-s2-s13
11. Kel, A., Boyarskikh, U., Stegmaier, P., Leskov, L.S., Sokolov, A.V., Yevshin, I., Mandrik, N., Stelmashenko, D., Koschmann, J., Kel-Margoulis, O. and Krull, M. Walking pathways with positive feedback loops reveal DNA methylation biomarkers of colorectal cancer. *BMC bioinformatics.* Cambridge (UK): RSC Publishing. **2019**;20(Suppl 4):119:1-20. doi:10.1186/s12859-019-2687-7
12. Filimonov D, Poroikov V. Probabilistic Approaches in Activity Prediction. Varnek A, Tropsha A. *Cheminformatics Approaches to Virtual Screening.* Cambridge (UK): RSC Publishing. **2008**;:182-216.
13. Filimonov DA, Poroikov VV. Prognosis of specters of biological activity of organic molecules. *Russian chemical journal.* **2006**;50(2):66-75 (russ)
14. Filimonov D, Poroikov V, Borodina Y, Gloriozova T. Chemical Similarity Assessment Through Multilevel Neighborhoods of Atoms: Definition and Comparison with the Other Descriptors. *ChemInform.* **1999**;39(4):666-670. doi:10.1002/chin.199940210

Thank you for using the Genome Enhancer!

In case of any questions please contact us at support@genexplain.com

Supplementary material

1. Supplementary table 1 - Up-regulated genes
2. Supplementary table 2 - Down-regulated genes
3. Supplementary table 3 - Detailed report. Composite modules and master regulators (up-regulated genes in Experiment: cisplatin-resistant vs. Control: cisplatin-sensitive).
4. Supplementary table 4 - Detailed report. Composite modules and master regulators (down-regulated genes in Experiment: cisplatin-resistant vs. Control: cisplatin-sensitive).
5. Supplementary table 5 - Detailed report. Pharmaceutical compounds and drug targets.

Disclaimer

Decisions regarding care and treatment of patients should be fully made by attending doctors. The predicted chemical compounds listed in the report are given only for doctor's consideration and they cannot be treated as prescribed medication. It is the physician's responsibility to independently decide whether any, none or all of the predicted compounds can be used solely or in combination for patient treatment purposes, taking into account all applicable information regarding FDA prescribing recommendations for any therapeutic and the patient's condition, including, but not limited to, the patient's and family's medical history, physical examinations, information from various diagnostic tests, and patient preferences in accordance with the current standard of care. Whether or not a particular patient will benefit from a selected therapy is based on many factors and can vary significantly.

The compounds predicted to be active against the identified drug targets in the report are not guaranteed to be active against any particular patient's condition. GeneXplain GmbH does not give any assurances or guarantees regarding the treatment information and conclusions given in the report. There is no guarantee that any third party will provide a refund for any of the treatment decisions made based on these results. None of the listed compounds was checked by Genome Enhancer for adverse side-effects or even toxic effects.

The analysis report contains information about chemical drug compounds, clinical trials and disease biomarkers retrieved from the HumanPSD™ database of gene-disease assignments maintained and exclusively distributed worldwide by geneXplain GmbH. The information contained in this database is collected from scientific literature and public clinical trials resources. It is updated to the best of geneXplain's knowledge however we do not guarantee completeness and reliability of this information leaving the final checkup and consideration of the predicted therapies to the medical doctor.

The scientific analysis underlying the Genome Enhancer report employs a complex analysis pipeline which uses geneXplain's proprietary Upstream Analysis approach, integrated with TRANSFAC® and TRANSPATH® databases maintained and exclusively distributed worldwide by geneXplain GmbH. The pipeline and the databases are updated to the best of geneXplain's knowledge and belief, however, geneXplain GmbH shall not give a warranty as to the characteristics or to the content and any of the results produced by Genome Enhancer. Moreover, any warranty concerning the completeness, up-to-dateness, correctness and usability of Genome Enhancer information and results produced by it, shall be excluded.

The results produced by Genome Enhancer, including the analysis report, severely depend on the quality of input data used for the analysis. It is the responsibility of Genome Enhancer users to check the input data quality and parameters used for running the Genome Enhancer pipeline.

Note that the text given in the report is not unique and can be fully or partially repeated in other Genome Enhancer analysis reports, including reports of other users. This should be considered when publishing any results or excerpts from the report. This restriction refers only to the general description of analysis methods used for generating the report. All data and graphics referring to the concrete set of input data, including lists of mutated genes, differentially expressed genes/proteins/metabolites, functional classifications, identified transcription factors and master regulators, constructed molecular networks, lists of chemical compounds and reconstructed model of molecular mechanisms of the studied pathology are unique in respect to the used input data set and Genome Enhancer pipeline parameters used for the current run.